

Wanted: social entrepreneurs

Scientists-turned-entrepreneurs are resuscitating the research and development of drugs for neglected diseases. Researchers, administrators and funders should contribute their expertise to help these initiatives — or set up their own.

The all-embracing open-source encyclopedia Wikipedia doesn't have a dedicated page on kala-azar, or visceral leishmaniasis. But who cares? After all, the disease only transforms vast numbers of people in developing countries into walking skeletons carrying bellies bloated by an enlarging liver and spleen. With drugs costing up to US\$200 a course, it often goes untreated, causing some 200,000 deaths each year. In the research and development (R&D) chains that lead to drugs, more attention is devoted to silicone breast implants and pills for erectile dysfunction than to the roughly 8,000 orphan diseases. These neglected diseases each touch up to just 2,000 people, but together they affect millions.

The root cause is that the markets for these diseases are too small to seriously interest the large pharmaceutical companies, which are involved in partnerships to address them but must also answer to their shareholders. What can be done to redress this balance by the many stakeholders along the route from research to product?

A round table on neglected diseases, co-organized by *Nature* at the BioVision World Life Sciences Forum in Lyon, France, last week, left a depressing sense of how far we have all got to go to seriously address neglected diseases — but also provided some encouraging sparks of enlightenment, and pointers to possible ways forward.

The meeting was held to celebrate the 50th anniversary of the approval of a polio vaccine. Polio is on the brink of being eradicated, something that would have been impossible without the March of Dimes, which courageously supported key basic research into the disease, and the Rotary Foundation, which got the vaccine into the field by raising more than \$500 million and providing volunteers. They addressed critical bottlenecks, and won. As a result, most people reading this will be unaware that just a few decades ago, polio was a disease that every parent feared.

New leaders

A new generation of 'social entrepreneurs' are testing imaginative ways to tackle neglected diseases. One small step for Victoria Hale, for example, may be a giant leap for mankind: the founding of a San Francisco-based not-for-profit drug company, the Institute for OneWorld Health. At the Lyon meeting she told of clinical-trial results that could lead to a safe and effective cure for kala-azar for just \$10 a course.

Instead of trying to develop a new drug from scratch, Hale started knocking on university and company doors, looking for drug leads lying in drawers, prescribed for something else or abandoned somewhere along the drug pipeline. She arm-twisted companies to give her promising candidates for such diseases as schistosomiasis and Chagas' disease, and went on to raise philanthropic funds for their clinical development.

The \$10 kala-azar drug is paromomycin, an off-patent antibiotic dating from the 1950s. With the help of funding from the Bill & Melinda Gates Foundation, Hale's company took it through the expensive clinical development and trials that no one else had ever bothered doing, to test and market it for kala-azar. Delivery will be achieved with the help of a generic-drug manufacturer in India.

A similar approach is being taken by the Geneva-based Drugs for

Neglected Diseases Initiative, instigated by Doctors without Borders, for a host of other neglected diseases. As well as testing and reformulating existing drugs, such initiatives are bridging the abyss between research and clinical development by focusing on leads that would otherwise have remained in the lab for lack of an industry sponsor.

Another key bottleneck in the pipeline is preclinical research in animal models or cell cultures. Genomics has revealed the molecular basis of some 1,200 orphan diseases, as well as major ones such as malaria and tuberculosis, revealing new drug targets. This is opening up many orphan diseases to possible cures for the first time. If the molecular basis involves the expression of interleukin A, for example, then existing interleukin A inhibitors might make potential drugs. But the many drug analogues needed by academic researchers to test this plethora of emerging targets are found only in industry.

So French gene therapist and social entrepreneur Alain Fischer is trying to persuade companies to give researchers access to the wealth of relevant drugs they hold but have not developed. The fledgling European Rare Disease Therapeutic Initiative gives the company the right of first refusal to market a promising molecule for a neglected disease. If it declines, the researchers can take it to another company or funder to take it to the clinic.

Persistence for change

The catch is that all these initiatives are relatively small. Funding is a big issue, but Hale is bullish. The money is out there, she says, provided that scientists, doctors and social entrepreneurs build awareness and argue their case. Robert Scott, who spearheads the Rotary campaign against polio, agrees, pointing out that governments, companies and research institutes change: someone who says 'no' today may say 'yes' tomorrow. Dogged persistence pays off.

Ironically, getting rights and molecules from universities is often harder than obtaining them from companies. University technology offices tend to patent aggressively, look no further than generating income, and often fail to include provisions beneficial to tackling orphan diseases in their licensing deals with companies. More scientists and their institutions should sign up to organizations such as the Centre for the Management of Intellectual Property in Health Research and Development, Biological Innovation for Open Society, and the Science Commons, which help academics to exploit their research and intellectual property for social ends.

The scale of the market failure in neglected-disease R&D is staggering, and awareness of the scale of the problem is chronically low among many research and health stakeholders. Sustained and inspired leadership is needed from those who shape research agendas or can bend the ear of funders, both individual and institutional, private and public.

Scientists, retired industry execs, fundraisers and others can volunteer research and expertise to the Institute for OneWorld Health (www.oneworldhealth.org/how) and other social entrepreneurs, or they can network at OrphanXchange (www.orphanxchange.org) and the US National Institutes of Health's Office of Rare Diseases (www.rarediseases.info.nih.gov). And, of course, they can highlight the challenge by adding the appropriate pages to Wikipedia. ■



Rising hopes

Will cardinals elect a Pope who backs use of condoms?
p944

Culture shock

Health workers struggle to contain Marburg disease
p946

**Broad scope**

NASA boss revives idea of manned flight to Hubble
p947

Back to basics

US drug regulator makes bid to reassert its authority
p948

Heightened security after flu scare sparks biosafety debate

Erika Check, Washington

US scientists who work with deadly strains of influenza are facing tighter restrictions on how they handle the viruses.

The rule changes follow last week's revelation that a 1957 pandemic flu strain was shipped to thousands of labs around the world in recent months.

Although biosafety experts have welcomed the measures, some argue that such accidents will only be prevented if there is a shift in attitude among those who work with dangerous viruses and bacteria.

On 12 April, the World Health Organization (WHO) announced that an H2N2 strain of influenza had been sent around the world as part of a routine quality-control test. It is now thought to have reached at least 6,000 labs across 19 countries. The flu samples were sent out under the auspices of the College of American Pathologists (CAP) based in Northfield, Illinois, which uses the kits to confirm labs' abilities to carry out diagnostic tests.

Usually, these kits contain flu strains that are not thought to pose a public health threat. But the H2N2 strain caused a worldwide pandemic in 1957, killing millions of people before it disappeared from widespread circulation in 1968. People born after this date have no immunity to the virus, raising concerns that it could reignite a pandemic if it was introduced back into the population.

The CAP says the company that produced the test kits, Meridian Bioscience of Cincinnati, Ohio, made the decision to include H2N2. But Meridian has not spoken to reporters, so it is unclear why it decided to use this strain.

"This was a very serious mistake," says flu virologist Robert Webster of St Jude Children's Research Hospital in Memphis, Tennessee. "This virus has killed millions of people. It should be locked away so that this never happens again."

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Pandemic flu viruses can now be handled only by biosafety level 3 labs.

On 18 April, a spokesman for the US Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, said that the agency was tightening restrictions on the types of labs that can work with strains of flu virus that have caused pandemics, including H2N2. These viruses are being upgraded from biosafety level 2 to biosafety level 3, a move that some other countries, including Canada, had made before last week's news.

Under wraps

The decision to designate pandemic flu viruses as level 3 will require labs that work with them to be specially constructed to restrict access and cut the risk of contaminants escaping. In addition, the CDC could recommend extra safeguards, such as requiring researchers to take decontamination showers every time they leave the lab.

The CDC has also said that it will review standards and practices for proficiency testing. And the CAP has promised to issue more exact instructions about how test kits should be constructed.

Federal regulation of pathogens in research in the United States is a highly contentious issue, especially since policy-makers have become concerned about the risk of a bioterrorist attack. The CDC tracks pathogens that are considered 'select agents'.

But the rules that govern such agents have met with resistance from scientists because they require labs to meet extensive security, registration and reporting requirements (see *Nature* **420**, 451; 2002).

The CDC and National Institutes of Health have joint power to decide what safety and security procedures should be followed by labs that handle pathogens that are not select agents, such as human flu viruses. But the new rules are unlikely to be popular with researchers.

Some argue that the whole affair has been blown out of proportion, and fear that changing the biosafety level of influenza could have a chilling effect on research, by restricting the number of labs that can work with the virus.

"I think it's overkill," says flu virologist Peter Palese of the Mount Sinai School of Medicine in New York. "Four thousand labs have received this virus and nothing has happened. It's really proved that the virus is not as dangerous as people make it out to be."

Biosecurity experts, meanwhile, support the changes — although some warn that if similar accidents are to be avoided in the future, scientists need to take biosafety more seriously. This would prevent incidents such as one last year, in which three researchers at Boston University developed tularaemia after they were exposed to the bacterium that causes it in a lab. And it would catch serious errors in judgment, such as the decision to send out a pandemic flu strain as part of a routine kit, says Stephanie Loranger, director of the Biosecurity Project at the Federation of American Scientists, an arms-control watchdog based in Washington, DC.

"The way we talk about biosafety is that it's something we have to do, not something we want to do," Loranger says. "I don't think there is the robust conversation about biosecurity in the scientific community that there needs to be."

Conclave kindles hope for bioethical reform

Declan Butler, Paris

Questions raised by medical advances and the spread of AIDS will create significant challenges for the next Pope, say religious experts.

Such considerations will be important factors in the deliberations of this week's electing conclave. But will a new leader create scope within the Roman Catholic Church for a shift in thinking on these issues? Few expect any rapid change, in particular on the Church's 135-year dogma that human life begins at conception, and its consequent opposition to human stem-cell and embryo research (see 'When does life begin?'). But AIDS workers are hopeful that there may be room for manoeuvre when it comes to the new Pope's position on condom use.

John Paul II was one of the most science-friendly Popes yet (see *Nature* 434, 684; 2005). But this openness ended where sexual issues came into play. Despite the fact that his pontificate spanned the start of the AIDS pandemic, the discovery of HIV and the spread of the disease throughout Africa, John Paul took an ultraconservative stance on the use of condoms. He embraced Pope Paul VI's controversial 1968 encyclical *Humanae Vitae*, which was the Church's first explicit ban on contraception.

On the encyclical's 20th anniversary, John Paul said that no consideration could justify the use of contraception, arguing that chastity and research are the answers to controlling the AIDS epidemic. His message has been amplified around the world: evangelicals in the United States have made fidelity



Catholic tastes: the Pope elected by this week's conclave may have more liberal views on condom use.

and abstinence prominent parts of President George W. Bush's African AIDS programmes, at the expense of efforts to promote condom use (see *Nature* 430, 279; 2004).

Many AIDS workers argue that John Paul's active opposition to condoms damaged efforts to control the disease, especially in poor countries. "Policy differences with the Vatican on condom use are an issue of concern," says Ben Plumley of the Joint UN Programme on HIV/AIDS in Geneva. He explains that the Church has hampered the agency's attempts to educate people. Advice from the Vatican has at times contradicted scientific understanding — one senior Vatican official declared on television that HIV is

small enough to pass through condoms (see *Nature Medicine* 9, 1443; 2003). But, he adds, "the Church in recent years has shown willingness to debate the issue. We strongly encourage this."

John Paul's position was by no means accepted throughout the Church. *Humanae Vitae's* stance on contraception has been controversial from the outset and is ignored by many of the clergy, according to Timothy Thibodeau, an expert on Church history at Nazareth College in Rochester, New York. "Many of the cardinals from the developing world are at the 'ground zero' of AIDS," he says, and they see the dogma as divorced from reality.

As John Paul's grip on the Church diminished in his dying days, opposition to his AIDS policy surfaced. In February, Cardinal Georges Cottier, the Swiss theologian of the pontifical household, became the most senior church official to argue that condom use to prevent AIDS is "legitimate". Cottier says a theological case could be made for using condoms to prevent the spread of disease, while forbidding their use as contraceptives. The same position was adopted earlier in February by the Spanish bishops' conference in Madrid. And a relaxed stance on condoms for AIDS prevention is also supported by Cardinal Carlo Martini, archbishop of Milan, Italy.

As *Nature* went to press, no new Pope had been chosen. However, Thibodeau warns that even a Pope with liberal views will not endorse change lightly. The problem is not contraception as such, he points out, but the fact that it is wrapped up in a 'package deal' of teaching on human sexuality. Many in the church fear that relaxing dogma on contraception could trigger wider questions on topics such as marriage for priests and the notion of sex for pleasure.

When does life begin?

The Roman Catholic Church has been a vigorous opponent of research on human embryos, including work on stem cells and therapeutic cloning. Despite lively debates, the next Pope is unlikely to change that.

The dogma that human life begins at the moment of conception was first introduced by Pope Pius IX in 1869. He put an end to centuries of thinking, inspired by Thomas Aquinas, which suggested that the embryo does not acquire a soul until later in its development.

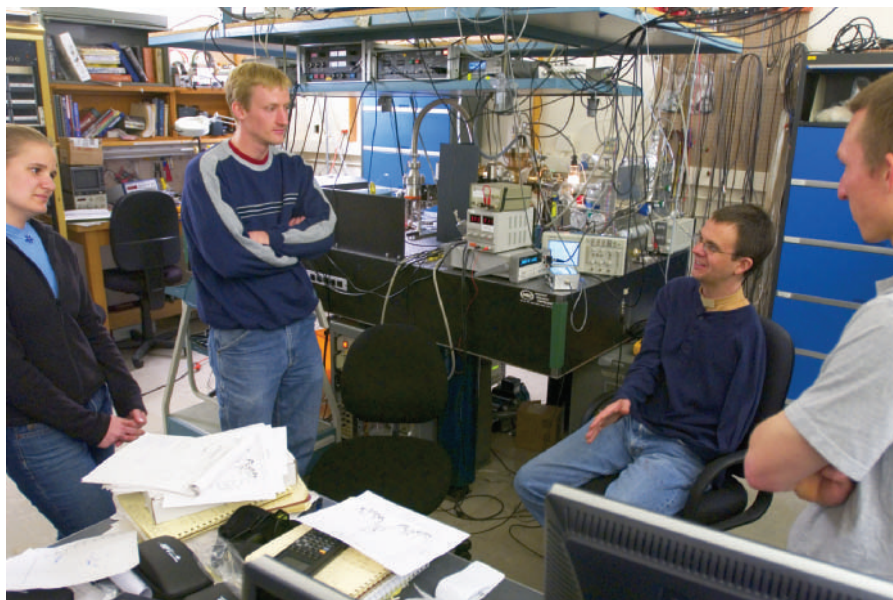
There is scientific evidence that the moment of conception takes several hours, and that twins and triplets can form up to 14 hours afterwards. This continues to fuel debates in the Church about when 'ensoulment' occurs. But Gregory Robbins, a historian of religion at the University of Denver, Colorado, says he does not expect any softening on the issue from the successor to Pope John Paul II.

"Ensoulment, including the provision of souls for what will become twins or triplets, presumably takes place within the purview of God's

foreknowledge," he says. "That does not seem to pose much of a theological problem. Embryonic and therapeutic stem-cell research do."

John Paul reiterated that life begins at fertilization in his 1995 *Evangelium Vitae*. The encyclical was based on an extensive scientific review of the literature commissioned from an international panel of scientists, says Roberto Colombo, who is head of genetics at the Catholic University of the Sacred Heart in Milan, Italy, and a member of the Pontifical Academy for Life — the Vatican's bioethics committee.

The Church concluded from the review, says Colombo, that although there are new hypotheses for markers of the beginning of a human — such as the first mitotic cell division or the appearance of an embryonic structure called the primitive streak — there is no consensus on these. The Church is therefore "resting on the solid ground of fertilization", Colombo says, adding that getting scientific consensus on alternative start points will probably take decades.



Much missed: Eric Cornell (sitting) is thrilled to start supervising his physics students again.

Nobel laureate triumphs over loss of arm and returns to lab

Kendall Powell, Boulder

Eric Cornell has experienced life's extremes. He won the physics Nobel prize in 2001, but last October he lost his left arm and shoulder — and nearly his life — to a flesh-killing bacterial infection. He has, however, bounced back and was in playful mood last week as he announced his return to the scientific fray.

"Losing an arm is more an inconvenience than a catastrophe," Cornell told a 12 April press conference organized by the University of Colorado at Boulder. "I should emphasize that I was right-handed before. And here it is." He waved at the assembled reporters.

Cornell will be fitted for a prosthetic arm, and hopes to have an attachment that will allow him to play pool. "The bets will go down, and then I'll put on the arm," he joked.

Doctors say Cornell is lucky to be alive after contracting necrotizing fasciitis, a rare and invasive streptococcal infection. Although he cannot recall injuring himself, the bacteria probably entered through a scratch or cut.

On 24 October 2004, he developed flu-like symptoms. The next day he felt a pain in his shoulder that steadily worsened, and within three days he was in the emergency room. Doctors surgically removed his left arm and shoulder and placed skin grafts on his torso. In a medically induced coma until late November, Cornell returned home from the hospital in mid-December.

Still undergoing physical therapy for his skin grafts, Cornell has phantom sensations of the missing arm: he feels as if it's behind his back. And he is taking medication for phantom pain, a common problem for amputees.

Cornell has already returned part-time to his lab at JILA, the institute run jointly by the

University of Colorado and the National Institute of Standards and Technology where he and his colleague Carl Weiman created a new form of matter called a Bose-Einstein condensate. That achievement won the duo the 2001 Nobel Prize in Physics, which they shared with Massachusetts Institute of Technology researcher Wolfgang Ketterle.

"Before I got sick, I thought I had one of the best jobs in the world," Cornell said. Thrilled to be back, he has been adjusting to using voice-recognition software instead of typing, and he won't be operating lathes or lasers. But he will continue to supervise students, which was taking up most of his time before his illness anyway. "It has to be said that it's been a little while since I've done very many fine adjustments on a laser," he said. "I usually offer helpful advice to my students and I'm as much a motormouth as ever."

That advice was much missed by lab members during his absence, says Aaron Leanhardt, a postdoc who arrived at Cornell's lab the week after he went into the hospital. "It will certainly be nice to be able to walk into his office and have a spontaneous discussion," Leanhardt says.

Cornell is now embarking on several studies involving Bose-Einstein condensates, as well as searching for an elusive property of the electron known as its permanent dipole moment. Physicists usually think of electrons as tiny, negatively charged spheres, but the charge might be asymmetrically distributed across the particle. That makes electrons a bit like people, Cornell told reporters: "Not one of us is really symmetrical."

His wife, Celeste Landry, couldn't resist adding a punchline: "You're one to talk!" ■

Additional reporting by Geoff Brumfiel.

Ecologist sues for lost tenure following transgene quarrel

Rex Dalton, San Diego

A Mexican ecologist renowned for his criticism of transgenic crops is suing the University of California, Berkeley, which denied him tenure in 2003.

Ignacio Chapela was prominent among staff at Berkeley who opposed a five-year, \$25-million deal that in 1998 gave the Swiss firm Novartis privileged access to findings by the university's plant scientists. The lawsuit claims that he was denied tenure in retaliation for this stance. Chapela also claims to be the victim of racial discrimination. "My case shows the tenure review process is totally overwhelmed by the forces of politics and the realities of economic dependency," he says.

Chapela rose to international attention with a 2001 paper reporting that transgenes had flowed into native varieties of maize in southern Mexico and fragmented throughout their genomes (D. Quist and I. H. Chapela *Nature* 414, 541–543; 2001). After the paper was criticized by plant biotechnologists and subjected to additional review, *Nature* issued a statement saying that it would not have been published had certain technical flaws been uncovered. But Chapela stands by the findings.

As these events were unfolding, Berkeley was considering Chapela's request for tenure in the College of Natural Resources. Although Chapela's department and the college voted in his favour, the university's chancellor at the time, Robert Berdahl, denied him tenure, saying his research record was insufficient. Since then, Chapela's case has become a focus for protesters concerned about threats to academic freedom from industrial forces.

Last year, a compromise was worked out in which university administrators agreed to convene a committee to reassess Chapela's tenure. This committee will make a recommendation to Berkeley's new chancellor, Robert Birgeneau, who was inaugurated on 15 April. He will rule on whether Chapela should receive tenure by 30 June — the end of the academic year.

Chapela says he is not optimistic about the review: "I see the academic process as unable to deal with the questions raised here." For now, he remains a salaried employee with an office and lab, but no funds for research.

The university declines to comment on the legal action. ■

Marburg workers battle to win trust of locals

Emma Marris

An outbreak of Marburg disease in Angola's Uíge province is spreading, despite the best efforts of several relief organizations. Western aid workers have been unable to win the trust or understanding of locals, and have even come under attack.

Marburg is a rare virus from the same family as Ebola. Few people survive the haemorrhagic fever it causes, but in general the virus is contracted only from people who

are visibly sick or dead. In theory, it should be simple to contain an outbreak by keeping such people in isolation.

"There is no vaccine, there is no cure. We have to break the transmission cycle by letting people know not to have contact with the dead," says Dave Daigle, a spokesman in Uíge for the World Health Organization. However, this causes conflict with families, who want to nurse their sick relatives and prepare bodies for burial.

Workers are trying to quarantine patients, and are sprinkling bleach on body bags in an effort to replace the strong local custom of washing dead bodies. But the paper-suited, respirator-wearing strangers who arrive at the homes of suspected cases by truck or helicopter and attempt to remove the ill or dead from their families have been met by resentment and, on several occasions, rocks.

Relief organizations are scrambling for ways to break this cultural impasse. Aid workers are now accompanied by anthropologists and health educators to help mollify families and convince them to part with their relatives. Religious leaders, including the local Catholic bishop, are being drafted to disseminate news about steering clear of the infected, and the Angolan Red Cross is going door to door.

The remaining members of a traditional band that lost its lead singer to the virus have even put out a song with lyrics about defeating the disease and cooperating with health workers. Called *Marburg*, it is being played on the radio and from the loudspeakers of roving trucks.

The World Health Organization and other aid agencies are working to contain the disease, but the death toll, now at more than 200, continues to rise. "I'd like to tell you that the numbers are contained, but they are not," says Daigle. ■

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Deadly contagion: aid agencies in Angola are working to quarantine patients dying of Marburg fever.

Computer conference welcomes gobbledegook paper

Philip Ball

Most graduate students would be delighted to have a paper accepted for presentation at an international scientific conference. But Jeremy Stribling, a computer-science graduate at the Massachusetts Institute of Technology (MIT) in Cambridge, wasn't sure whether to be amused or alarmed.

His paper, "Rooter: a methodology for the typical unification of access points and redundancy", co-authored with Daniel Aguayo and Maxwell Krohn, was accepted for the 9th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI), to be held in Florida, in July. But Stribling didn't write it; he let a computer do it.

Stribling and his colleagues have developed an 'automatic computer-science paper generator' that cobbles together articles adorned with randomly generated graphs. The 'results' are totally spurious.

The MIT researchers say they hoped to cause "maximum amusement" by aping the jargon of the less illustrious papers in computer science. But they also had a more serious goal: to test whether such

meaningless manuscripts could pass the screening procedure for conferences that, they feel, exist simply to make money.

'Rooter' passed the test: the WMSCI accepted it, albeit without peer review. The paper claims, among other things, that "the famous ubiquitous algorithm for the exploration of robots by Sato *et al.* runs in $\Omega((n + \log n))$ time".

It's not the first example of a hoax paper aimed at exposing poor reviewing and meaningless jargon. In 1996, US physicist Alan Sokal published a paper on the "hermeneutics of quantum gravity" in the journal *Social Text*. Sokal's paper parodied the post-modernist language of some contributions to that publication and prompted a vigorous debate about the intellectual respectability of 'cultural studies'.

The WMSCI conferences have been running for ten years, and last year's meeting attracted nearly 3,000 papers. WMSCI 2005 advertises itself as "trying to bridge analytically with synthetically oriented efforts, convergent with divergent thinkers".

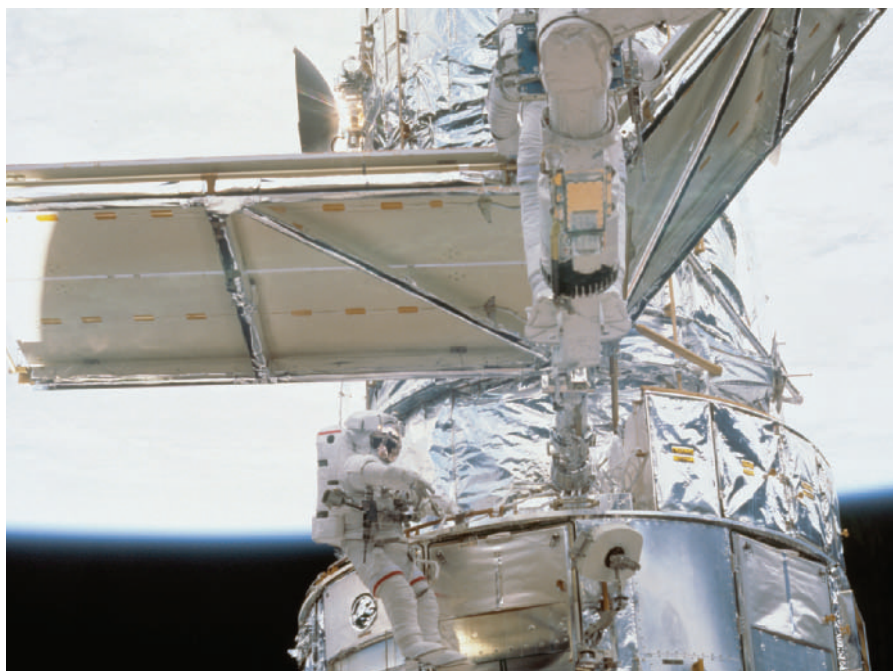
The MIT team regards it as one of many

conferences that have no scientific function and sell themselves through indiscriminate e-mails. "You see lists of speakers, and there's no one you've ever heard of," says Stribling. "They spam us."

Such conferences have sparked anger in the field, as demonstrated by a WMSCI submission from David Mazières of New York University and Eddie Kohler of the University of California, Los Angeles. The title, text and figures of their ten-page paper consist entirely of the phrase "Get me off your fucking mailing list".

"I don't know why these conferences exist," adds Frans Kaashoek, a member of the MIT computer-science group to which Stribling and his colleagues belong.

But the WMSCI's general chairman, Nagib Callaos, who is based in Venezuela and has no listed academic affiliation, has defended the conference's decision. "We did not receive reviews for some papers," Callaos says. "Since we thought that it was not fair to reject those, we accepted them as non-reviewed ones." The MIT paper has now been pulled. ■



Back in the picture: astronauts could once more be sent to service the Hubble Space Telescope.

NASA boss takes the helm and launches plans for the future

Tony Reichhardt, Washington

NASA's new administrator, Michael Griffin, got off to a flying start last week before he had even been sworn in. At his Senate confirmation hearing on 12 April, he reopened the possibility of sending astronauts to service the Hubble Space Telescope, an option his predecessor Sean O'Keefe ruled out last year (see *Nature* **427**, 273; 2004).

The next day, the Senate approved Griffin as the space agency's 11th — and perhaps most technically savvy — leader. Members of Congress had nothing but praise for the 55-year-old engineer and physicist, who arrives at NASA at what he calls a “watershed moment”, as the United States tries to reinvigorate human spaceflight with an ambitious plan to send astronauts beyond Earth orbit for the first time in decades.

Griffin wholeheartedly supports that goal, and he's impatient to get started. NASA had planned to have a new Crew Exploration Vehicle ready for outbound astronauts by 2014, but Griffin told senators that he hopes to speed up the pace of development. “People want a space programme that goes somewhere and does something,” he told NASA employees on 14 April.

First, though, he'll focus on getting the shuttle flying again — before retiring it as soon as the International Space Station is completed “in a manner consistent with our international partner commitments”.

O'Keefe's decision to scrap the idea of sending a shuttle to service the ageing Hubble telescope was made in the immediate

aftermath of the doomed Columbia shuttle mission, Griffin said, adding that he wants to revisit the plan once the shuttle has successfully flown again — a mission planned for late next month (see *Nature* **434**, 811; 2004).

Griffin also called for a balanced programme of science, astronaut exploration and aeronautics. He claimed that NASA can afford them all, pointing out that the agency's current budget is equivalent to that at the time of the Apollo Moon missions in the 1960s and 1970s.

He also defended NASA's science programme, calling it “one of our crown jewels”. “I don't agree that it has fallen by the wayside,” he added, but admitted that he could understand why scientists are worried. Spending on the new Moon–Mars programme and the shuttle's return to flight “has caused temporary dislocations in science funding”, he said, that are likely to continue for the time being.

In his exchanges with senators and NASA employees, Griffin was as direct as his predecessor had been discursive, showing how he earned himself a reputation as a straight talker.

Before the congressional committee, he referred to himself as “a simple aerospace engineer from a small town”. Standing before his new employees, he stressed humility and collegiality. “I'm not sir, I'm not Dr Griffin. I'm Mike or Michael,” he said before the packed auditorium at the agency's Washington headquarters. “The NASA administrator is not royalty.”

Scientists speak out in search of fame and fortune

Jim Giles, London

Three minutes, no slides and immediate, hard-hitting feedback from a panel of judges: tough conditions for even the most seasoned of orators. But for the past few weeks, hundreds of UK scientists have been subjecting themselves to this ordeal in a bid to win the science-communication competition FameLab.

The regional heat in London on 14 April saw contestants wax lyrical about everything from the science behind the “perfect design of the penis” to the difficulty of defining what is and isn't a drug. Each then faced the enthusiastic but occasionally blunt thoughts of the judging panel, while their audience watched.

Many entrants say that they see FameLab as a way of developing outreach skills. “Science communication will be a big part of my career,” says Ed Moran, who is studying for a PhD in infectious diseases at the University of Oxford. His talk involved a bedpan containing chocolate spread, and discussed how human faeces could be used to replenish gut bacteria.

But it was Matt Wilkinson, a zoologist at the University of Cambridge, and Rebecca Lloyd-Evans, a biotechnology consultant with Cambridge-based firm BioBridge, who impressed the six judges enough to secure places in the 12-strong final, to be held in June at the Cheltenham Festival of Science. The pair triumphed with respective talks on animal locomotion and the link between manic depression and creativity.

The overall winner will get £2,000 (US\$3,800), a part in a TV science programme on Britain's Channel 4 and a speaking tour. The competition, organized by the Cheltenham festival and the UK National Endowment for Science, Technology and the Arts, was inspired by the hit TV show *Pop Idol*, in which aspiring pop stars audition before a celebrity panel.

The judges say that the quality of entries has been high. “I've been really impressed,” says Kathy Sykes, who holds a chair in science communication at the University of Bristol. She notes that scientists often have more passion for their subject than professional communicators — and that passion and enthusiasm, rather than a polished performance, were priorities for the judges.

“Everyone has been really good,” adds science author and judge Simon Singh. “We haven't had the *Pop Idol* thing where people just can't sing.”

www.famelab.org

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Big issue: the FDA is considering allowing the cosmetic use of some silicone breast implants.

Painkillers taken off the shelf as FDA acts on staff warnings

Washington The embattled US Food and Drug Administration (FDA) is taking tough decisions over products in two highly profitable areas of medicine.

Congress has attacked the FDA over the past year, after some of the agency's staff criticized its ability to monitor drug safety (see *Nature* **434**, 554–556; 2005). Regulation of one class of drug singled out for criticism — the blockbuster painkillers known as COX-2 inhibitors — has now been tightened. On 7 April, the FDA asked drug

manufacturer Pfizer to withdraw its Bextra painkiller from the market.

The FDA is also considering an expert-panel report on silicone breast implants. The panel said on 13 April that implants made by Mentor of Santa Barbara, California, could be used in breast-augmentation operations but advised against the use of a similar product made by local rival Inamed. Concerns about ruptures and other complications led the FDA to effectively ban silicone implants for cosmetic use in 1992.

Robot craft arrows to target but suddenly darts away

Washington A US spacecraft designed to dock with satellites without human intervention missed its planned rendezvous last week.

The Demonstration of Autonomous Rendezvous Technology (DART) craft, part of a US\$110-million mission launched last Friday, was supposed to dock with a military communications satellite almost 800 kilometres above Earth. DART successfully tracked its target and came within 100 metres, but an unknown mishap caused it to divert its orbit away from the satellite. The craft was designed to manoeuvre itself to within 5 metres of the satellite.

NASA has announced that it will launch an investigation into the reasons for the mission's partial failure.

Asteroid warnings set to take milder tone

Washington Astronomers are to use less threatening language to describe asteroid risks in a bid to avoid scaring the public.

On the Torino Scale, zero equates to no danger, and ten to a certain collision and global catastrophe. But newly discovered asteroids often look threatening and generate excited media coverage, before being downgraded once their orbits are better understood. Last December, for example, an object briefly merited the first-recorded four but was reassigned to zero within days.

Worried that the scale was scaring people, its creators have worked up a revised version, unveiled on 12 April. Instead of “meriting concern”, objects rating two to four now only require “attention by astronomers”, say Richard Binzel of the Massachusetts Institute of Technology and his colleagues.

A second scale, the Palermo Technical Impact Hazard Scale, was developed in 2002 solely for use by astronomers. This scale allows researchers to decide which new sightings need to be tracked by telescope (see *Nature* **418**, 468; 2002).

Europe tightens controls on transgenic corn imports

Munich A \$400-million annual trade in US animal feed is under threat following the discovery of an unauthorized transgenic strain of corn, or maize, on American farms.

The European Union said on 15 April that it will require imports of corn-based feed to be certified as free of the unapproved variety *Bt10*. Last month, *Nature* revealed that US farmers had grown 15,000 hectares of *Bt10* corn by mistake between 2001 and 2004 (see *Nature* 434, 548; 2005). The US Department of Agriculture has fined Syngenta, the company responsible, \$375,000 for the error.

Bt10 contains a gene from the soil bacterium *Bacillus thuringiensis* (*Bt*) that produces a pesticide. The *Bt* gene has been approved for use in other varieties, but *Bt10* also contains a gene for antibiotic resistance. Syngenta immediately pledged to implement testing of animal-feed exports at major US ports, to enable certification, "within days".

Roman researchers face year on the road

Munich Eighty Italian cognitive scientists face a nomadic existence, with two laboratory moves in the space of three months, after being told to leave their

Fishing-line weights net conservation prize

San Diego Three technologies designed to reduce the environmental impact of fishing are due to scoop US\$35,000 in prize money from the conservation organization WWF.

A system of fishing-line weights, which forces an entire long-line system below a depth of 100 metres, so sea turtles (pictured) won't be snagged on hooks and die, was due to be awarded first prize when the results of the Smart Gear competition are announced on 21 April in Washington. The system will scoop a \$25,000 prize for Steve Beverly, a fisheries development officer with the Secretariat of the Pacific Community in Noumea, New Caledonia.

One \$5,000 runner-up award was expected to go to a team that chemically treats gill nets with barium sulphate to make them more

current facility before a new home is built.

Researchers at the Institute of Cognitive Sciences and Technologies in Rome have been told that their building is too expensive to maintain and does not meet safety standards. The CNR, Italy's main basic-research organization and the institute's funder, said on 14 April that they have until the end of the month to leave. A permanent home for the institute on the outskirts of Rome is due to be complete

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

acoustically detectable, to help drive away whales, dolphins and porpoises. The second runner-up place was set to go to a team that devised a system to catch large shrimp in the bottom of nets while allowing juveniles to escape from the top.

by the end of the year, but is likely to be delayed by conflicts about who should foot the bill.

In the meantime, the researchers will lodge at another CNR facility in Rome for three to four months before moving to as yet undefined locations. "Our research will be endangered by having to constantly move our lab around the city," says institute scientist Rino Falcone. The CNR is not making any statements on the decision.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

CSI: cell biology

Digital photography and image-manipulation software allow biologists to tweak their data as never before. But there's a fine line between acceptable enhancements and scientific misconduct. Helen Pearson investigates.

In a cramped office in midtown Manhattan, a forensics expert peers intently at a flickering computer screen. The shadowy image, hugely magnified, reveals a tell-tale dark smear. Something about it, she can tell, is just not right...

It could come straight from the screenplay of the latest hit TV crime show. But, in fact, such scenes are playing out regularly at the sedate headquarters of Rockefeller University Press — where the images under scrutiny are those of cells and gels in papers accepted for publication in *The Journal of Cell Biology*.

Mike Rossner, the journal's managing editor, introduced the forensics procedure in 2002 to patrol a growing practice among cell and molecular biologists: that of manipulating their scientific images. His specially trained editor hunts for tell-tale lines or smudges that might reveal where inappropriate modifications have been made — and when found, Rossner asks the authors for their original data.

In the vast majority of cases, the perpetrators aren't willfully misrepresenting their results — rather, they are unaware that their efforts to achieve the cleanest images for publication have crossed the line of acceptability. Such ignorance is widespread, and for Rossner it underlines why his journal now subjects all accepted papers to image forensics. "My goal is to catch problems before we publish," he says.

The availability of digital cameras and image-manipulation programs such as Photoshop has made it all too easy for researchers to enhance the DNA bands on a gel or brighten up the images of cells snapped on a slide. Most alterations are harmless: researchers legitimately crop a picture or enhance a faint, fluorescently tagged protein. But in some cases, innocent attempts to clarify an image can erase valuable data or raise suspicions of fabrication.

Following Rossner's lead, other journal editors now scrutinize images more carefully

than they used to. And researchers, some burnt by questions about their own photos' integrity, are beginning to patrol the activities of their lab members. "We all underestimate the amount of skulduggery that goes on," says Joseph Gall, who studies the structure of cell nuclei at the Carnegie Institution in Baltimore, Maryland.

Picture perfect

Biologists have always gone to great lengths to create beautiful pictures that best illustrate their data. Geneticists carry out multiple exposures of radioactively labelled DNA fragments separated on gels in order to create a crisp image. Cell biologists once worked through rolls of film to grab the perfect shot from a microscope. But with Photoshop, a few clicks of the mouse can transform a featureless black microscope snap into a starry vista littered with labelled proteins. "You can make up almost any image you want nowadays," says Tom Misteli, a cell biologist at the National Cancer Institute in Bethesda, Maryland.

Some biologists seem to succumb to this temptation. In 1989–90, only 2.5% of allegations examined by the US Office of Research Integrity, which monitors misconduct in biomedical research, involved contested scientific images. By 2001, this figure had jumped to nearly 26% (ref. 1).

But such miscreants are in the minority. The more pressing problem, say experts, is that innocent efforts to smarten or prettify images end up with unintended consequences. At the very least, biologists risk erasing potentially valuable information, such as

low levels of fluorescently labelled protein swilling around a cell's cytoplasm. At worst, such manipulations can lead researchers to the wrong scientific conclusions.

In some cases, researchers may not even realize that they are significantly altering an image, particularly when the changes are made at the time a picture is taken, by adjusting settings such as the exposure on microscopes or digital cameras. Many scientists are oblivious to the consequences of such actions, because they have only a rudimentary knowledge of the sophisticated equipment involved.

But it is the use of Photoshop or similar software to alter an original picture that scientists and editors say is most troublesome. With these programs, researchers can quickly carry out an array of modifications, from slicing off the messy edges of an image to using sophisticated algorithms to sharpen the edges of a blurred image on a microscope slide.

No one wants to ban image manipulation outright. In cell-biology experiments, for example, researchers often have to adjust the relative intensities of red, green and blue fluorescent markers in order to show all three in a single image. Even drastic changes are sometimes considered tolerable if scientists spell out exactly what they did.

But it is tough to draw a precise line between acceptable and unacceptable image manipulation. Few journals have explicit policies, and of those that do, *The Journal of Cell Biology* has the most stringent guidelines². These allow alterations that are applied equally across an entire image, such as changes to contrast or brightness. They also permit some other corrections, such as adjusting the brightness of pixels in a certain range of colours — but only if details of the adjustment are spelled out. Changes to selected parts of an image, such as brightening one cell in an entire field or scrubbing out an ugly blemish, are prohibited.

Blurred vision

Rossner estimates that roughly 20% of accepted manuscripts contain at least one figure that has to be remade because of inappropriate image manipulation. In the vast majority of cases, the authors have made the changes innocently, and can provide acceptable images once the problem has been explained. In rare cases where the journal suspects misconduct, the authors' institutions are informed.

For their part, scientists say that they feel under pressure to produce faultless images to present convincing experiments that reviewers and editors want to publish. "The temptation comes from the fact that you have to sell a clear-cut story," Misteli says. Tweaking images is also seductive in a way that adjusting statistics is not, because of the natural human desire to create an aesthetically pleasing picture.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Caught on camera: image enhancements designed to clarify fluorescently labelled cell structures can also wipe out useful data.

Cell-signalling researcher Shigemi Matsuyama is one biologist who discovered the pitfalls of Photoshop when the images in a 2003 *Nature Cell Biology* paper³ were called into question. His study showed how a protein called Bax, which is involved in triggering cell death, is controlled by a second protein called Ku70. At the time, Matsuyama was excited about the publication, which was his first after setting up his lab at the Blood Research Institute in Milwaukee, Wisconsin.

But his elation turned to despair when a colleague in the field contacted him and *Nature Cell Biology's* editors a few months later to point out some troubling features in several images. The pictures were of western blots, a technique in which a labelled antibody is washed over several protein samples and bands on the blot to reveal which of the proteins bind to the antibody. Enlarging several of these images revealed straight lines on either side of some bands, suggesting that they had been pasted together on a computer. "I was stunned," Matsuyama says. "I couldn't eat for almost a week."

The first author of the paper, postdoctoral researcher Motoshi Sawada, had used



In the frame: Mike Rossner has championed the fight against unwarranted image manipulation.

Photoshop to put together some of the western blot images. In one example, he cut out a band from one position on the blot and pasted it into a second spot. In another, he reused bands from one blot in two separate figures. Sawada says that he felt under pressure to produce the figures quickly, and that one of the changes was a genuine mistake. He adds that he made the others so that the figures were clear and easy to understand. "I thought that it was acceptable," he says.

In the event, Sawada and Matsuyama went back to the original data, and were able to produce corrected figures that confirmed their original scientific conclusions. These were published in a 2004 corrigendum⁴.

Image problem

Sawada's case seems to be fairly typical. Scientists and journal editors say that most questionable image manipulation can be traced to inexperienced students or lab staff who are unclear about what is allowable. "It's junior people tidying up the image and not realizing that what they're doing is wrong," says Richard Sever, executive editor for the *Journal of Cell Science*, based in Cambridge, UK.

But whoever the perpetrators are, and whatever their motivation may be, there is so far little consensus about how to curtail them. One way, researchers say, would be to teach budding biologists about the ethics of image-making during postgraduate courses. Another is for lab heads to be more rigorous about policing inexperienced lab members. Matsuyama, for one, now demands that his students and postdocs show him the original image alongside the final form that they want to submit for publication.

Journal editors are also recognizing that they have a role to play. The *Journal of Cell Science* plans to draw up image manipulation guidelines for its authors within the next three months; *Nature Cell Biology* is encouraging researchers to submit original images of gels alongside the edited ones, for publication as supplementary information⁵. Another option would be for journals to demand that authors list the image adjustments they make in methods sections or figure legends.

Despite their heightened level of scrutiny, editors still fret that their forensic efforts will be outpaced by rapid advances in image-manipulation software, and researchers' skill at using it — the minority of biologists who are deliberately bending the rules are only likely to get better at covering their tracks. "We're just fortunate that most students and postdocs are not that good in Photoshop yet," Rossner says.

Helen Pearson reports for news@nature.com from New York.

1. Krueger, J. W. *Accountabil. Res. Pol. Qual. Assur.* **9**, 105–125 (2002).
2. Rossner, M. & Yamada, K. M. *J. Cell Biol.* **166**, 11–15 (2004).
3. Sawada, M. et al. *Nature Cell Biol.* **5**, 320–329 (2003).
4. Sawada, M. & Matsuyama, S. *Nature Cell Biol.* **6**, 373–374 (2004).
5. *Nature Cell Biol.* **6**, 275 (2004).



IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

The mating game: having used a robotic bowerbird (left) to examine courtship behaviour, researchers plan to use the same technique on sage grouse (above).

When robots go wild

A steady stream of mechanical animals is marching out of the lab into the field. Jonathan Knight tunes in to see how these motorized models can expose what makes real creatures behave the way they do.

As the sage grouse gather on the Wyoming prairie to find mates this spring, a few of the strutting, posturing males are in for a surprise. One of the hens they'll be trying to court will be a mechanized impostor rolling around on tracks from a model railroad.

A dozen metres away, crouching in the brush, behavioural biologist Gail Patricelli will be guiding its every move by remote control—ready to rescue her decoy should things get out of hand. “The worst-case scenarios are the bird derails and gets torn up by the males, or an eagle swoops in and grabs it,” says Patricelli. Either would put an abrupt end to an experiment she has spent months setting up.

Patricelli, an assistant professor at the University of California, Davis, is one of several technically savvy biologists who have latched on to the use of robotics in behavioural studies. Radio transmitters, computer chips, digital cameras and audio recorders have become smaller and cheaper, placing home-made robots within the reach of even junior researchers on a tight budget.

As a result, a menagerie of robots—from squirrels to lobsters—has been deployed to test ideas about animal behaviour that had previously been too tricky to tackle. Robots are making new avenues of research possible, says behaviouralist Christopher Evans at Macquarie University in Sydney, Australia. “We tend to seize on technologies like this and say: ‘I’ve been waiting to do that experi-

ment for 20 years, and now I can,’” he says.

Among the first to explore the potential of mechanized animals was Axel Michelsen at the University of Southern Denmark in Odense. In the early 1990s, he built a mechanically controlled honeybee of wax-coated brass to help him decode the insects’ dances, which tell others in the hive where to find food. The brass bee was slightly larger than normal and had just one wing, fashioned from a broken razor blade. It didn’t look much like a real bee. Fortunately, beehives are completely dark.

Because they can’t see the dancer, honeybees perceive its movements mainly through the pulsing currents of air it creates. By making his bee perform sections of the dance to an audience of real bees, Michelsen discovered how the dance revealed both the direction to and the distance from food¹.

Wired for romance

Patricelli was so fascinated by this approach that she built a robot bird for her postdoctoral research on the courtship of satin bowerbirds (*Ptilonorhynchus violaceus*). It was fairly simple: a stuffed female that could crouch, turn its head and fluff its feathers at the tweak of a joystick. To real males it was utterly convincing. Patricelli found that males modified the intensity of their display in response to cues from the female: hamming it up if she seemed receptive and toning it down if she was aloof².

The bowerbird has been retired to a card-

board box while Patricelli works on the more sophisticated grouse-bot. This radio-controlled model not only turns and makes provocative crouching movements, it is also equipped with video and audio recorders to catch the males’ responses.

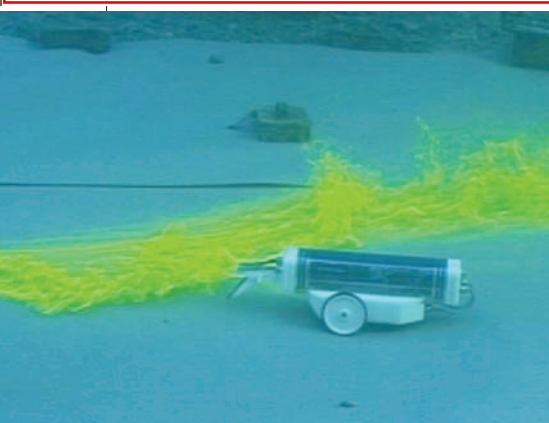
The point of the experiment is to monitor the grouse’s mating signals. “It’s very hard to tease apart the conversation when you are just watching from the outside. But when you control one side of it, you get a better idea of what is going on,” Patricelli says. Sage grouse (*Centrocercus urophasianus*) mate in leks, open fields where males strut around, displaying their feathers and emitting a loud drumming noise from vocal sacs on their chests. Like real females, the robot will wander through the lek, ‘observing’ different males and showing interest by moving its hindquarters. Patricelli hopes to learn which of the hen’s moves are most exciting to a male and how he responds.

Giving a robot natural movement is fairly easy for birds, as their motion is somewhat jerky and robotic to begin with. In fact, the wheeling motion of Patricelli’s grouse may be a bit smooth, but that won’t necessarily matter, as animals seem to respond to very specific cues when identifying members of their own species. For instance, Evans and his Macquarie colleague Ann Göth last year built a series of robot chicks to find out what cues newly hatched brush turkeys (*Alectura lathami*) look for in spotting their nest mates. All

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS



Fur will fly: with a tail that heats up and swings from side to side, this robotic rodent (above) is helping to test theories about how ground squirrels try to frighten away real rattlesnakes (above left).



Not to be sniffed at: a robotic lobster attempts to follow a model odour plume to its prey.

the robots looked the same, but it was the one that made the jerky pecking motion typical of game birds that got all the attention³.

Robots are also proving their worth in studying communication between different species. Just across campus from Patricelli is a room housing several dozen rattlesnakes (*Crotalus viridis*) that, courtesy of graduate student Aaron Rundus, face regular encounters with a robot squirrel.

California ground squirrels (*Spermophilus beecheyi*) are a favourite prey of rattlesnakes, but they are not always easy targets. Adults are resistant to snake venom and so can be quite bold, often driving off a snake with an aggressive display that involves swinging their tail like a windscreen wiper.

What is unique about this display is that it seems to use both visual and infrared cues. Rattlesnakes are exquisitely sensitive to heat, and a squirrel will heat up its tail by as much as 2 °C during tail-flagging. As Rundus told the annual meeting of the Animal Behavior Society in Oaxaca, Mexico, last June, the same squirrel seems to keep its tail cool when flagging a heat-insensitive gopher snake (*Pituophis melanoleucus*).

To test how much the rattlers' response is

influenced by the heat signal, Rundus teamed up with Davis mechanical engineer Sanjay Joshi to rig a stuffed squirrel with a motorized, heat-controlled tail. Rundus is coy about discussing his unpublished findings, but hints that things are going well. What's important, Rundus says, is that robotics has made possible a controlled experiment that couldn't have been done in the wild — there is no way to convince a squirrel to keep its tail cool when facing a rattler. "This is long overdue in behavioural biology," he says.

Mock lobster

Researchers can learn a lot by building robots that mimic animal behaviour, even if the machines don't interact with live creatures. For instance, Frank Grasso, director of the Biomimetic and Cognitive Robotics Lab at the City University of New York in Brooklyn, has studied how lobsters sniff out their prey by using a robo-lobster that follows a scented plume in the water. "The robots allowed us to improve on ideas that had been around for a long time," he says.

It had been assumed that lobsters locate their prey by swimming towards higher concentrations of odour. But in lab trials, a robo-lobster working on this principle alone did poorly, particularly if it was a long way from its target. Guessing that this might be the result of turbulence muddying the odour gradient in the water, Grasso's team gave the robot a sense of flow, allowing it to assume that odours come from upstream. The robot then behaved much more like the real thing⁴.

But even so it didn't quite act naturally. Grasso has continued to refine the model, for instance by adding a memory algorithm to help it if it loses the odour plume. "This is allowing us to build step by step a model of what is going on in the lobster's brain," he says.

The approach has its limits. Even if the

real and fake lobsters behave the same, there is no direct test to show that the robot algorithm and a lobster's brain are making decisions in the same way. But such experiments can at least establish a minimum, says psychologist Jeffrey Schank, who is working with Joshi to build robotic rats that also follow simple behavioural rules. "If we can get robots to mimic the rats, that's the first step to understanding them," Joshi explains.

The robo-rats are entirely autonomous, and make decisions about where to go based on which sensors are activated in combination with a probability algorithm programmed into their computer brains⁵. For Schank the important question is: how little does the brain have to do? "We are trying to quantify the lowest level of cognitive ability necessary for certain types of behaviour," he says. This can then feed into theories of how animals behave in the wild.

Certainly, the robotic toolbox has only just been prised open. In addition to one-on-one interactions, some researchers are looking at group behaviour. A European team led by Jean-Louis Deneubourg at the Free University of Brussels has built robo-roaches that can interact with groups of live cockroaches. The researchers are using them to understand collective decision-making, such as choosing a safe hiding place.

The herd instinct may also encourage more behavioural researchers to use robots. "Every time I go to a conference, I hear about more people using robots," says Patricelli. "It's catching on."

Jonathan Knight writes for *Nature* from San Francisco.

1. Michelsen, A., Andersen, B. B., Storm, J., Kirchner, W. H. & Lindauer, M. *Behav. Ecol. Sociobiol.* **30**, 143–150 (1992).
2. Patricelli, G. L., Uy, J. A. C., Walsh, G. & Borgia, G. *Nature* **415**, 279–280 (2002).
3. Göth, A. & Evans, C. S. *J. Exp. Biol.* **207**, 2199–2208 (2004).
4. Grasso, F. W. *Biol. Bull.* **200**, 160–168 (2001).
5. Schank, J. C., May, C. J., Tran, J. T. & Joshi, S. S. *Adap. Behav.* **12**, 161–173 (2004).

J. McDONALD/CORBIS (SNAKE); A. RUNDUS

Who's helping to bring science to the people?

With student numbers falling, we need more researchers to do public-outreach work.

Sir — Your Commentary article “Weapon of mass attraction” (*Nature* **433**, 357–358; 2005) highlighted poor participation by US scientists in public-outreach activities.

It may be interesting to look at the situation in France, where the reform of the national research agency, the CNRS, has led to the creation of a working group on the popularization of science, which I lead.

Preliminary data on activities carried out by the 10,403 CNRS scientists, from July 2003 to June 2004, show that three-quarters of them did not get involved in any popularization, or public-outreach, activity during the year — taking this to mean writing a book or presenting a popular-science lecture or poster.

A closer look at the data reveals that efforts are unequally distributed: the most active 10% of scientists account for 70% of all public-outreach activity and the top 5% account for half.

A fit of the histogram of the number of activities per scientist produced a more precise picture. We discovered that a single Poisson distribution cannot fit all the data. This confirms that scientists do not constitute a homogeneous population.

The best fit we found requires three different Poisson distributions, which can be interpreted as three different

sub-populations: the ‘silent majority’ (76% of scientists, who do hardly anything); the ‘open minority’ (21%, who carry out some activity once or twice a year); and the ‘semi-professional popularizers’ (3% of scientists who carry out activities on average six times a year). Researchers in this active minority dedicate a significant fraction of their research time to the public, accounting for a third of all activities in bringing science to the people.

Policies aimed at getting scientists more involved in public outreach should perhaps be tailored to the three sub-populations we found. For example, those who do not yet carry out any public-outreach activities have to be convinced of the importance of doing so. The decline in student numbers seems to have persuaded more physicists to get involved in the World Year of Physics. The ‘open minority’ might be encouraged by access to simple tools for efficient public outreach, while the ‘semi-professional popularizers’ may have concerns about institutional rewards.

Contrary to expectations, our data suggest that age is not significant. We find that the average number of public-outreach activities is broadly constant with age, increasing moderately as scientists get older, from 0.45 activities a year for

those aged 31–35, to 0.7 a year at 56–60.

Our data also reveal variation among fields: the proportion of scientists carrying out public-outreach activities varies from 17% for general physics, chemistry and biology, to 30% for astrophysics and 41% for social sciences.

Interestingly, while the mean number of activities per scientist varies considerably from field to field (from 1.2 for social sciences to 0.3 for general physics, chemistry and biology), the productivity of active scientists is constant across all fields (close to 2.5 actions a year).

Finally, our data reveal that speaking at conferences on popular science is the most common activity (25%), followed by writing newspaper articles (23%) and giving radio or television presentations (17%).

Again, there are variations across different disciplines. For example, much higher numbers of social scientists appear on radio or television (61%). So any researcher willing to pursue television might be able to learn something from their colleagues in the humanities.

Pablo Jensen

Laboratoire de Physique de la Matière Condensée et Nanostructures,
Université Lyon I,
69622 Villeurbanne Cédex, France

Reviewers not attached to online submission

Sir — John P. Moore’s criticism, in Correspondence, of electronic journal submission, “Online submission makes authors do all the work” (*Nature* **433**, 800; 2005), only examines the author–editor interaction. The poor peer reviewer also suffers from electronic manuscript submission.

I am generally happy to review papers relevant to my own expertise, unless I’m too busy. But papers forwarded as e-mail attachments almost invariably cause headaches when I am unable to open or print some or all of the associated illustrations. It recently took two days to find someone with the necessary know-how and software to open and print a plate at publication size, yet it took me less than an hour to review the associated paper.

Given the choice, I always ask editors to send me a hard copy. I am never so free as to be able to review a paper immediately, so the delay of even intercontinental mail is not an inconvenience.

Worse still are some grant-awarding

bodies. Recently, I couldn’t review a proposal when asked by the US National Science Foundation. But just saying “No” by return e-mail is not acceptable. Instead, it is necessary to open the attachment and follow instructions to the “Decline” box.

Steve Donovan

Department of Palaeontology,
Nationaal Natuurhistorisch Museum, Postbus 9517,
NL-2300 RA Leiden, The Netherlands

Funding shouldn’t rely on competing death tolls

Sir — I think it is important to note the irony in your News story, “Protest letter accuses health agency of biodefence bias” (*Nature* **434**, 7; 2005), of researchers who shun the study of disease complaining that a US agency called the National Institutes of Health fails to fund them adequately.

I find it striking that those who protest against the funding of biodefence research are proposing instead that public-health menaces should be given the highest priority. By this standard, many of the letter’s signatories should voluntarily

return their funding for research on *Bacillus subtilis*, *Escherichia coli* and other non-pathogens so that it can be appropriately directed towards the obvious public-health threats of HIV and tuberculosis.

Although I disagree with their reasoning, I strongly agree that basic microbial research is so important, in and of itself, that its funding level should be not be reduced regardless of other concerns, whether arising from public health or bioterrorism. Using body counts (“Bio-weapons agents cause, on average, zero deaths per year”) may be useful in the short term to frame the debate, but I fear they will be damaging in the long run. How many of us want to be asked, when our next grant is reviewed: “How many people did your bug kill last year?” I certainly don’t.

If basic research is relevant to the health of the nation, then make the case that it is so. The current approach will only leave funding levels vulnerable to the next media sensation or hysterical distraction.

David Hilbert

Department of Anatomy and Cell Biology,
P&S 12-517, Columbia University, New York,
New York 10032, USA

Seeds of discord

A useful, though partial, survey of how we breed the plants we eat.

Mendel in the Kitchen: A Scientist's View of Genetically Modified Foods

by Nina Fedoroff & Nancy Marie Brown
Joseph Henry Press: 2004. 370 pp. \$24.95,
£17.99 (hbk)

Margaret E. Smith

Most people in the industrialized world are blissfully ignorant about plant breeding. Plants provide the food in their supermarkets, flowers to grow in their gardens, and raw materials for industries that produce many items they find useful. Yet people are unaware of plant breeders' efforts over many years to improve the yield and quality of cultivated plants using a growing suite of genetic knowledge and technologies.

The advent of genetically engineered crop plants has, if nothing else, made a certain sector of the public — those who object to such innovations — intensely aware of the discipline of plant breeding. In *Mendel in the Kitchen*, Nina Fedoroff and Nancy Marie Brown set out to address a number of concerns that have been raised about this most recent addition to the plant breeder's tool-box.

The book provides an informative and engaging description of the history of agricultural science in general, and of plant breeding in particular. The origin of the ambiguous concept of the 'species' is well explained, for instance, highlighting the difficulty of defining the 'traditional species barrier'. Before genetic engineering was possible, cross-breeding was largely limited to plants that are sexually cross-compatible — a criterion that was used to define the limits of a species. Now genetic engineering allows genes to be moved between organisms, regardless of cross-compatibility, so it has been criticized for breaking the species barrier.

The history of other widely accepted plant-breeding tools helps to put genetic engineering in a realistic context. 'Wide crosses' — those between a crop and its wild relatives — are nothing new, nor is the technique of chromosome doubling: the grain triticale, for example, is a product of a wide cross between wheat and rye, followed by doubling the chromosome number to make it fertile. Mutagenesis — using chemical or radiation treatments to cause genetic changes — may sound questionable to some but has been used to produce a few crop varieties that are grown and eaten without concern. Our current agricultural and horticultural plants are far from 'natural' and have been genetically modified by humans

Healthy debate: golden rice may alleviate vitamin A deficiency, but will it reduce rice diversity?

for thousands of years using a broad range of tools, some of them more drastic in their effects than others.

The facts presented in Fedoroff and Brown's book are, for the most part, accurate. It is the things they choose not to include, and the inclusion of some sweeping generalizations, that give the book its decidedly pro-genetic-engineering slant. The omissions may have been meant to minimize detail and complexity, but their accumulated effect is to give the book a distinct bias that people worried about this technology will readily detect.

For example, in a discussion of 'golden rice', which has been genetically engineered to contain beta-carotene in an effort to alleviate vitamin A deficiency, the authors say that concerns about restriction of the gene pool are unfounded. If golden rice proves popular, there is no reason to expect that only one strain will be grown, they argue, as the trait could be bred into any of the thousands of rice varieties grown by farmers at present. This is technically true. However, golden rice is intended to be given to subsistence farmers free of charge, so there is no economic incentive for the private sector to move the beta-carotene genes into thousands of varieties. Public-

sector breeding programmes, meanwhile, have neither the personnel nor the financial resources to do so. Initially only a few golden rice varieties will be developed: varieties that have appropriate adaptation, sufficient yields, and acceptable eating quality for the major rice-growing environments. If golden rice is to have the impact on vitamin A deficiency that its proponents claim, it will need to be very widely grown. That will happen only if these few varieties replace a diversity of existing ones, just as the original 'green revolution' rice varieties did a few decades ago. We now have a sense of the value of that earlier loss of diversity, and should aim to avoid repeating that history.

The authors' approach to the StarLink incident is also enlightening. StarLink corn carries a particular variant of the *Bacillus thuringiensis* (Bt) toxin that the US Environmental Protection Agency (EPA) initially approved for use in animal feed but not as food, pending further research to clarify results that suggested that the Bt toxin might have allergenic potential. The authors document this well, along with the subsequent recalls and financial settlements that resulted when StarLink was found in food. But their implication that the EPA, in imposing the restrictions, was somehow responsible for

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

the losses incurred by Aventis, the company marketing the StarLink corn, is absurd. Both Aventis and the EPA can be faulted for pretending that corn grown for animal feed could be kept completely separate from corn grown for human consumption. But Aventis chose to market StarLink as animal feed only, rather than waiting to collect the additional data required by the EPA for approval as a food. The issue is not whether data ultimately showed any human food risk from StarLink (in fact, there appears to be no such risk). It is one of public confidence in the regulatory effectiveness of the EPA and in the company's compliance with the restrictions under which it agreed to operate.

Chapter 7 of the book boldly states that we now know that "recombinant DNA technology is among the safest technologies ever developed" — a broad and sweeping claim for a technology that is only a few decades old. In his recent review of Michael Crichton's novel *State of Fear* (*Nature* 433, 198; 2005), Myles Allen noted: "A hallmark of good science must be the way it treats uncertainty." In this, Fedoroff and Brown have not been as forthcoming as they should. As a plant breeder, I fully understand the frustration of scientists who are focused on the good that a tool such as genetic engineering can do. However, the agricultural and plant-breeding history that Fedoroff and Brown describe, with its theories that later proved untrue and its technologies that were harmful in unanticipated ways, suggests that a degree of humility would be appropriate. Just as opponents of genetic engineering are unaware of, or are loath to acknowledge, the aspects of this technology that we do understand, such as the genetic history of our cultivated plants, so proponents are reluctant to admit the ambiguities and unknowns about genetic engineering.

As a well-written, engaging account of a controversial subject from a scientist's viewpoint, *Mendel in the Kitchen* should be on the reading list for everyone interested in genetic engineering, both proponents and opponents. It assembles a large and informative body of information about many of the issues that have raised concerns about genetically engineered crops. However, although the authors state in their introductory chapter, "Which view will seem right to you depends on what you consider conventional, and on how you define the ways of nature," the rest of the book attempts to convince readers that only one view is right. ■

Margaret E. Smith is in the Department of Plant Breeding and Genetics, Cornell University, Ithaca, New York 14853, USA.

More on food and genes

Genes on the Menu: Facts for Knowledge-Based Decisions

by Paul Pechan & Gert E. de Vries
Springer, £23, \$39.95, €29.95

Admitting sympathy beyond species

Thinking with Animals: New Perspectives on Anthropomorphism

edited by Lorraine Daston & Gregg Mitman
Columbia University Press: 2005. 240 pp.
\$49.50, £32

Juliet Clutton-Brock

When Konrad Lorenz began his studies of animal behaviour in the 1930s, thereby marking the beginnings of ethology as a science, he related the ways of his jackdaws to the ways of people and used highly personal and anthropocentric language. Then, from around the middle of the twentieth century, with the onset of behaviourism, any assumption that animals share the thoughts, feelings or motivations of humans began to be called anthropomorphism. It became a pejorative term applied to the ideas of anyone who dared to believe that animals were capable of conscious thought.

Over the past decades, with the expansion of knowledge about the behaviour of humans and animals, as well as an increased understanding of evolutionary theory, psychologists and philosophers have joined ethologists in research into the minds of animals. This research, known as cognitive ethology, has resulted in many investigations into consciousness, cognition, self-awareness and intelligence, as well as on whether animals feel pain, anger, fear or love, or have a theory of mind. In this context, the title of philosopher Thomas Nagel's essay "What is it like to be a bat?" (in *Mortal Questions*, Cambridge University Press, 1979) has been much quoted and paraphrased in the ethological literature, including this book.

With cognitive ethology has come the general realization that anthropomorphism does not necessarily disrupt scientific observation but can support the continuity between humans and animals. A strong supporter of this view is Frans de Waal, who proposes the term 'anthropodenial' for the rejection of shared characteristics between humans and animals (*The Ape and the Sushi Master*, Basic Books, 2001).

On the other hand, there are still those who maintain that anthropomorphism is a distraction from scientific rigour. Over half a century, J. S. Kennedy never wavered from his published view that "if the study of animal behaviour is to mature as a science, the process of liberation from the delusions of anthropomorphism must go on" (*The New Anthropomorphism*, Cambridge University Press, 1992).

All these points are cited and discussed in the nine eclectic essays in *Thinking with*

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Rotating photographs of bats turns our perceptions about them upside-down.

Animals, many of which begin with a sentence such as "Anthropomorphism has long been considered a bad word in science". However, as the somewhat contentious title suggests, the main theme of the book is not a further discussion of the now rather jaded arguments for or against the humanizing of animals in science. Instead, it examines the actual practice of anthropomorphism in a wide variety of case studies.

One of the most impressive of these shows how taking an anthropocentric attitude to animals and treating them as individuals can have a dramatic effect on attitudes to animal conservation. In a fascinating account, Gregg Mitman outlines the management of elephants in Kenya. In the 1960s, when 'hard' science ruled, thousands of elephants were killed in order to keep the population within what was considered to be the carrying capacity of Kenya's national parks and reserves. Then, in the 1970s, came the work of Iain Douglas-Hamilton, followed by many others, to make the public aware of the social life of elephant families. This new perspective led directly to the promotion of the elephant as an endangered species and the banning of the ivory trade.

An essay by the film-maker Sarita Siegel describes how she made a documentary on the work of Anne Russon with orang-utan orphans. This demonstrates the importance for conservation of cooperation between the media and biologists with attitudes

Science and culture

Polluting utopia

The idyllic reputation of a lost island community may not be wholly deserved.

Andy Meharg

The remotest occupied island in the British Isles, the Outer Hebridean outlier Hirta in the St Kilda archipelago, has become symbolic of a utopia corrupted by modern, capitalist society. Originally a self-sufficient and self-governing community, the islanders changed their aspirations as soon as they had regular contact with outside culture in the nineteenth century. They yearned to leave their sea-girt, lonely home, and Hirta was completely abandoned in 1930. The death of the island community was immortalized in Michael Powell's 1937 film *The Edge of the World*.

The island has long commanded the fascination of the British public. It was the subject of one of the most remarkable, and earliest, of British travelogues, Martin Martin's *A Description of the Western Islands of Scotland*, circa 1695, and his later *A Voyage to St Kilda*. Martin travelled at the behest of the Royal Society, recording Gaelic culture before the disruption that followed the Jacobite rebellion of 1745. Unlike most early documenters of remote cultures, Martin was of the people, a Gael from the Hebridean island of Skye. His book inspired, among others, Samuel Johnson's *A Journey to the Western Isles of Scotland*.

The extent of anthropological study of the small St Kildan community is unparalleled, and has spawned a vast literature in the past 300 years. The documenters of island life were captivated by a culture solely dependent on the exploitation of seabirds, which were harvested for food, oil and feathers. They were also entranced by the island's remoteness and apparent independence. The archipelago's owner extracted a yearly rent. The island was self-sufficient and had no cash economy. It had a 'parliament' that met every day to take democratic decisions affecting the life of the community. A rotation of land stewardship ensured fairness in resource allocation, like the "anarcho-syndicalist commune" of *Monty Python and the Holy Grail*. Many nineteenth- and twentieth-century chroniclers have described the island as a utopia, including John MacCulloch, who wrote of his 1815 visit: "If this island is not the Eutopia so long thought, where will it be found? Where is the land which has neither arms, money, law, physic, politics, nor taxes?"



Paradise lost: the inhabitants of Hirta contaminated their island by overexploiting seabirds.

The Hirta community is now an iconic image for a devolved Scotland with a strong socialist tradition. The architect Enric Miralles's design for the new Scottish parliament drew inspiration from the St Kildan parliament, incorporating the forms of upturned keels of boats as a representation of the island's life.

But this utopia is a myth. The St Kildans' bird culture had severe effects on European seabird populations. Most notably, the islanders beat to death the last observed great auk in Britain, which they held to be a witch, in 1840, just before the global extinction of the species. My own research has thrown some light on the ecological balance of the island, showing that agronomic practices led to severe pollution of the arable land.

According to Martin, the St Kildans recycled all their wastes — faeces, peat ash, and bird guts and bones — applying them to their arable land as fertilizer. When I analysed these arable soils they were highly contaminated with lead, zinc and dioxins (*Nature* **421**, 909–910; 2003). By analysing the lead isotopic composition of the arable soil, archaeological and modern seabird bones, and ancient peat ash from dwelling floors, I worked out that this pollution was coming from peat ash

and bird bones in the compost stream. This pollution of arable soils is, to date, unique in terms of ancient settlements and was included in the case made to UNESCO (the United Nations Educational, Scientific and Cultural Organization) by the National Trust for Scotland (the archipelago's current owner) for extending St Kilda's World Heritage Site status for natural and marine environments to include the cultural landscape. The UNESCO committee will consider the proposal in June. Pollutant levels are so high, with lead and zinc concentrations in arable soils reaching 500 mg per kg, that they compare with the soils of Britain's most polluted industrial cities.

The pollution arose because this small island, with only 40 acres of arable land, has supported a relatively large population of around 200 inhabitants at least since the Iron Age. The island could sustain such a population only because the inhabitants acquired their protein from massive exploitation of seabird colonies. The accretion of human waste streams over the millennia has resulted in the polluted soil. Very far from paradise. *Andy Meharg is in the School of Biological Sciences, University of Aberdeen, Cruickshank Building, Aberdeen AB24 3UU, UK.*

that, although they are anthropomorphic, are based on sound science.

The other seven essays are on widely different topics, from Wendy Doniger on zoomorphism in Indian Sanskrit texts to Paul White on the experimental animal in Victorian Britain. Cheryce Kramer discusses the ways in which Tim Flach manipulates his photographs to change our perceptions of animals: for example, his photos of bats are printed so that the animals appear to be

standing upright, rather than hanging down, so they look like standing people. Sandra Mitchell asks: "What is the fate of anthropomorphism in contemporary science?" and argues that evidence is difficult to obtain, either way, for whether there are cognitive similarities between humans and animals. James Serpell writes about how thinking of animal companions in human terms is responsible for many of the benefits that their owners derive from them, such as the

apparent love and loyalty of dogs. For the dogs, however, anthropomorphic intervention such as the docking of tails cannot be condoned, he argues.

Thinking with Animals is an unusual book that will surely join the growing literature on consciousness, animal cognition and the continuity between human and animal minds.

Juliet Clutton-Brock, South Barn, High Street, Fen Ditton, Cambridgeshire CB5 8ST, UK.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Reaching the peak: what
are we going to do when
we run out of oil?

The descent of Mount Petroleum

Beyond Oil: The View from Hubbert's Peak

by Kenneth S. Deffeyes

Hill and Wang: 2005. 202 pp. \$24

Robert K. Kaufmann

Nostradamus has nothing on M. King Hubbert. In the late 1950s, Hubbert correctly forecast that US oil production, which was then increasing rapidly, would peak in 1970 and decline thereafter. But Hubbert's accuracy has generated a conundrum: repeated announcements of the peak in worldwide oil production made by Hubbert's acolytes gather considerable attention, but their perennial failure provides ammunition for the 'cornucopians', who argue that large supplies of oil remain. This is a false dichotomy, but Kevin Deffeyes encourages it. He frames the debate about world oil supplies as one between Hubbertists and cornucopians, and as such, *Beyond Oil* can be viewed as the latest salvo between two caricatures of the real oil situation. Nonetheless, there is valuable information to be gleaned from this thoughtful and entertaining book.

The first three chapters provide an excellent description of the historical context and mathematical basis of Hubbert's methodology. Unfortunately, Deffeyes ignores subsequent research which indicates that Hubbert was lucky as well as being a genius. Had prices evolved on some other path, or had the Texas Railroad Commission — the US prototype of OPEC (the Organization of the Petroleum Exporting Countries) — controlled production using some other criterion than price stability, Hubbert's prediction would probably have been less accurate.

Because of subsequent changes in these

economic and institutional forces, Deffeyes' updated version of Hubbert's analysis states that 228 billion barrels will be recovered from US oil fields, which is about 30% higher than Hubbert's estimate. Deffeyes does not mention this increase, perhaps because he views Hubbert as his 'patron saint', but by doing so he misses an opportunity to demonstrate the strength of Hubbert's method, which is that relatively large uncertainties about recoverable oil supply have relatively little effect on the timing of the peak in production. Deffeyes' contention that global oil production will peak in 2005 may be wrong (again), but that does not mean that the cornucopians are correct. Given most reasonable estimates for world oil supply, the peak is not far off, and that is why society must now consider what comes after oil.

This consideration is where the book shines. In the following six chapters, Deffeyes combines his expertise in science in general and geology in particular with anecdotes from his extensive work in the private sector as a petroleum geologist to assess potential alternatives to oil. Most are geological in origin, although he also devotes a chapter to hydrogen. For each, he describes its geological origins and the engineering difficulties that are associated with its production. In addition, he offers many engineering rules of thumb that influence the economic viability of the alternatives. Better than any economic cost-benefit analysis, this combination of theory and practical know-how gives the reader a better feel as to why there is still no obvious alternative to oil. As such, these chapters are ideal for non-specialists and undergraduates who are convinced that the world will eventually need a replacement for oil but wonder why people shy away from this need every time the price of oil drops from its most recent peak.

Ironically, Deffeyes denigrates under-

graduates in the subject where his book may be most appreciated: environmental studies. He comments several times that environmental-science courses are flawed because they fail to teach students 'hard science' and instead emphasize breadth. This may be true, but Deffeyes' focus on depth leads to some 'breadth-taking' errors of his own. Perhaps most startling is his argument that "the major oil companies are not saying publicly that the oil game is over. [But] if there were attractive prospects available, companies would be clawing their way over one another to get to the drilling rights." Deffeyes says they are conspicuously not doing so, but he fails to realize that many OPEC nations, especially those in the Middle East, prohibit foreign oil companies from investing directly in their oil fields. And yes, oil companies do claw their way over one another each time Saudi Arabia, Kuwait or some other OPEC nation opens a possibility for direct investment. Geologically attractive sites remain, but they are unavailable politically.

The interplay among geology, economics and politics is what makes the world oil market so interesting. No book can hope to do justice to all its aspects. Deffeyes' argument has some flaws, but no more than other books that try to describe the current energy situation and possible solutions for it. Despite these limitations, readers should remember that the book uses a methodology that generated a remarkably accurate forecast for one of the most important economic events in the twentieth century. And that creates the book's real take-home message: even if production does not peak in 2005, we all need to know that Hubbert's peak is coming soon, and we had better start thinking about a future without oil.

Robert K. Kaufmann is in the Department of Geography and the Center for Energy and Environmental Studies, Boston University, Boston, Massachusetts 02215, USA.

Striking a chord

Landscapes: when perturbed by climatic and tectonic changes, landscapes resonate with a range of frequencies.

Philip Allen

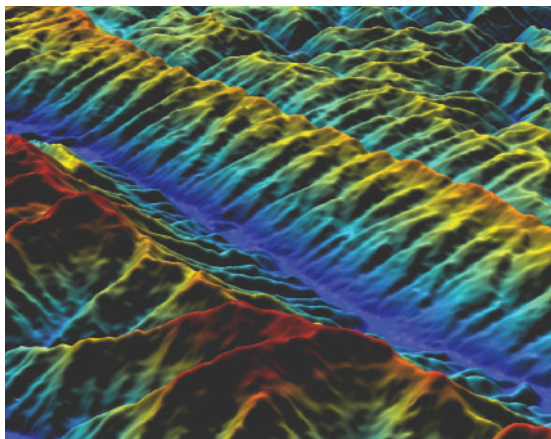
How is the physical form of a landscape linked to the wide array of processes that shape it? Such processes range from the patter of tiny raindrops, and the shuffling of beetles and rodents on hill slopes, to major slope failures that generate colossal debris flows, and cratering caused by the impacts of meteorites. To some, the issue is encapsulated in the conflicting imagery of 'catastrophism' and 'continuity' (or 'uniformitarianism') — two opposing views in a debate that is almost as old as the discipline of geology itself.

To take an analogy from the world of art, the uniformitarian or continuity school of thought sees the geological and geomorphological canvas as being made up of the gradual accumulation of minute brush strokes — much like that produced when a pointillist, such as Georges Seurat, adds the unremarkable dots to his picture. Landscapes viewed in this way are orderly, predictable, reassuring, perhaps even benevolent.

But another kind of canvas, such as a van Gogh, is characterized by wild brush strokes of vivid colour, and is restless, unpredictable and disconcerting. This catastrophic view sees landscapes, and indeed Earth history, as being shaped predominantly by exceptional events. If this is the case, present-day processes cannot fully explain the physical nature of the world around us: most large-scale events have not been experienced by humans, and can only be surmised from geological studies of 'deep time'.

The irony of the catastrophism versus continuity debate is that no real dividing line exists between the 'ordinary' and the 'exceptional'. For example, plots of data from thousands of landslides that occurred during the past 60 years in the Southern Alps of New Zealand describe a beautiful power law, supporting the view that landscape change follows a continuum in magnitude and frequency. This implies that it is impossible to separate the exceptional from the ordinary, and invites us to realign our thinking and develop a new vocabulary when investigating the processes that shape landscapes.

We can consider landscapes as the changing geometrical form of a critical interface between two interacting systems: an internal system driven essentially by tectonic fluxes of rock, and an external system dominated by the effects of climate. The challenging



Making waves: a digital reconstruction of a landscape from Kentucky, United States, reveals regular transverse catchments.

feature of these two interacting systems is that they each operate at a range of temporal and spatial scales. We can think of landscapes as being perturbed by variations in both internal and external mechanisms. The results of these perturbations might be seen as changes in the morphometric properties of the critical interface or as changes in the mass fluxes of rock, particulate sediment or solutes through and over it.

In the past, tectonic movements were somewhat statically viewed as effectively instantaneous events causing uplift of the land surface, and upon which geomorphic agents worked over eons of time. It is now clear that in regions of active tectonics, the rates at which tectonic and geomorphic processes operate are similar. Consequently, the possibilities for complex coupling of the two processes are enormous.

An exciting avenue of research at present is evaluating the response of landscapes to perturbations, such as changes in tectonic and climatic conditions. For example, the landscapes of the Basin and Range province in western United States are continually adjusting to the tectonic deformation of the underlying crust. This tectonic deformation is essentially caused by the growth of faults to accommodate extension of the crust, and is responsible for uplifted ranges with steep transverse rivers that feed cone-shaped 'fans' of sediment in adjoining basins. Although such a system appears, at first, to be relatively simple, its operation is determined by the timescales over which several interacting geomorphic subsystems respond to perturbations. Rather than resonating like a single hand-held bell, landscapes are polychromatic, resonating with different

frequencies when perturbed, like the strings of a guitar that generate chords when struck at different fret positions.

So instead of being seen as continuous or catastrophic, I propose that landscapes be considered in terms of their response times to various perturbations, such as the lateral growth of an active fault, or the change in the rate at which a major fault slips due to the accumulated effect of earthquakes. Landscapes can be viewed as 'buffered' when their response to rapidly repeating perturbations is slow, but 'reactive' when their response is rapid compared to the frequency of the disturbance. Landscapes can also be termed 'steady'

when their response time is smaller than the interval between a repeated perturbation, or less than the time passed since a change in the prevailing tectonic or climatic conditions, but 'transient' when the reverse is true.

One of the goals of the burgeoning field of research working on Earth surface processes must be to improve measurements of the rates of tectonic, sedimentologic and geomorphologic processes using the rapidly improving tools that are available. Such tools include the digital description of landscapes; the dating of geomorphic surfaces and events; the elucidation of the spatial patterns of denudation at different timescales; and the use of numerical methods in coupled geomorphic tectonic–climate modelling of integrated systems routing sediment from erosional mountain slopes to depositional basins. With these new data, we will be better able to establish what meaningful information about the forces shaping landscapes can be interpreted from its geometrical form, mass fluxes and sedimentary deposits. Put simply, we may learn how to read the epic poem of Earth history that is written in sediments and landscapes. ■

Philip Allen is in the Department of Earth Sciences, ETH-Zentrum, CH-8092 Zurich, Switzerland.

FURTHER READING

Summerfield, M. A. (ed.) *Geomorphology and Global Tectonics* (Wiley, 2000).

Allen, P. A. & Allen, J. R. *Basin Analysis: Principles and Applications* 2nd edn Ch. 7, 221–265 (Blackwell, 2005).

This essay is based on the William Smith Lecture given at the Geological Society of London's conference 'Earth's Dynamic Surface: Catastrophe and Continuity in Landscape Evolution', 4–5 October 2004.

Cracking the Himalaya

Douglas W. Burbank

The collision of India with Asia causes large earthquakes and active faults along the southern margin of the Himalaya. But has localized erosion by monsoon rains created new faults in the interior of the range?

The provocative idea that climatically driven erosion could govern tectonic deformation has instigated a decade of geodynamic models^{1–3} and geological studies^{4,5} that explore potential climate–tectonic feedbacks. According to this idea, erosion of mass from Earth's surface may determine where tectonic deformation is most rapid. Consequently, heavy precipitation, rapid erosion and active faulting are predicted to be spatially correlated in an active mountain belt, or orogen. Based on local variations in erosion within the Himalayan range, Wobus *et al.* (page 1008 of this issue⁶) deduce the presence of a large, previously undocumented fault that ruptures the surface, and which is interpreted as a response to especially intense rainfall.

One challenge in testing connections between climate and tectonics is that active faults are difficult to locate in the bedrock core of mountain belts: erosion commonly removes features, such as displaced river terraces, that record readily recognizable offsets. Although different types of bedrock may be juxtaposed across a fault, that in itself reveals little about how rapidly the fault slipped, or whether it last ruptured 100 million years ago or a decade ago. As a consequence, few active faults have been identified within the core of mountain ranges, even when it is known that rapid tectonic contraction is occurring across the range.

Wobus *et al.*⁶ use an innovative combination of techniques to deduce that there is a major surface-breaking fault within the interior of the Himalayan range. Rather than examine observable offsets of the surface, they use two contrasting measures of erosion rates to demonstrate an abrupt change in rates across a narrow zone. Their approach involves measuring the concentrations of cosmogenic radionuclides and the ratio between argon isotopes ($^{40}\text{Ar}/^{39}\text{Ar}$) in sediments, which are interpreted to record variations in erosion rates on thousand-year and million-year timescales, respectively. An argon 'cooling age' measures the time since cooling below around 350 °C for muscovite (the mineral dated by Wobus and colleagues), thereby defining an average cooling rate. When divided by a geothermal gradient (temperature change with depth), a cooling rate yields a mean erosion rate.

In mountain ranges where contraction



Figure 1 Blue yonder. Steep hill-slopes in the rapidly eroding Himalaya north of the newly defined fault⁶ and the 'physiographic transition'.

A. M. HEIMSATH

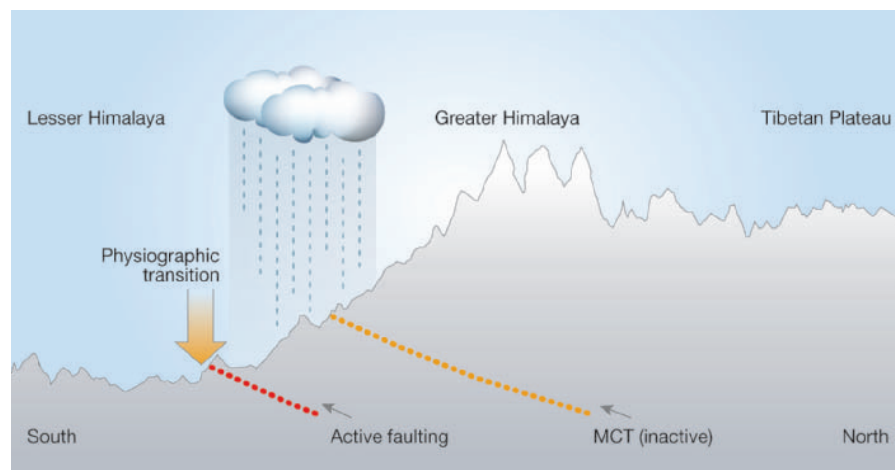


Figure 2 Topography of the Himalaya at the site of Wobus and colleagues' study⁶. The site lies in the northern Lesser Himalaya, at the physiographic transition just south of the Main Central Thrust (MCT), which marks the boundary between the Lesser and Greater Himalaya. Wobus *et al.* argue that the transition is a consequence of an active fault. The existence of the fault is inferred from the estimated rapid surface erosion, caused by heavy monsoon rainfall, which allows southerly 'channel flow' of the lower crust to reach the surface.

results from a long-lived continent-to-continent collision, erosion rates are commonly considered as proxies for rates of rock uplift. Consequently, the differential uplift implied by the spatially abrupt change in erosion rates, as identified by Wobus *et al.*, suggests that active faulting has persisted in the core of the Himalaya for millions of years. Moreover, noting that the highest monsoon rainfall occurs near the inferred fault, Wobus *et al.* link fault slip to climatically driven erosion, thereby coupling tectonics with climate.

Throughout the Himalaya, the boundary between two major subdivisions of the range, the Lesser and Greater Himalaya, is traditionally defined by the Main Central Thrust. This is a major, northward-dipping fault along which deeply buried rocks in the north have been thrust southwards over less deeply buried rocks. The fault described by Wobus *et al.* lies in the northern Lesser Himalaya, 5–20 km south of the Main Central Thrust, at a topographic change (termed the 'physiographic transition') where broad valleys and gentle hill-slopes are replaced by narrow valleys, steeper rivers, higher mountains and steeper hill-slopes (Figs 1 and 2).

The significance of an active fault in this position can be viewed in the context of three perspectives on collisional orogens. First, the traditional view is that, in such orogens, deformation migrates outwards over time, leaving more interior faults as relics, while creating faults at the distal edge of the orogen⁷. Most geologists consider the Main Central Thrust to have been inactive for the past 10 million years, but its footwall (the northernmost Lesser Himalaya) may have been undergoing deformation between 6 million and 2 million years ago^{8,9}. Wobus and colleagues'

results would extend this activity to the present.

Second, collisional orogens have been likened to tapered wedges¹⁰, in which the outward slope has a constant angle, even as the orogen grows, much like the wedge of snow in front of a snowplough. If erosion differentially removes material from the wedge's interior, contraction within the wedge is predicted to restore the original taper. The correlation of active faulting with active erosion on steep Himalayan slopes is consistent with this model.

Third, crustal thickening resulting from the Indo-Asian collision is predicted to have caused partial melting of the lower crust beneath the Tibetan Plateau¹¹. Driven by the high potential-energy gradient along the margins of Tibet¹², outward flow of the lower crust ('channel flow') is predicted to emerge at the surface as a fault-bounded zone where erosion is focused and vigorous². Wobus *et al.* suggest that their newly identified fault bounds the lower margin of the channel and that the erosion caused by locally high monsoon rainfall has concentrated faulting near the foot of the Greater Himalaya (Fig. 2). If so, this is a significant discovery of a zone of deformation that is intimately tied to climatic gradients.

These interpretations are not unequivocal. A more complex structural geometry could juxtapose the contrasting argon cooling ages found by Wobus *et al.*, without surface faulting. In the context of the 'snowplough' model, for example, a steeper step in the sliding surface beneath the wedge could migrate southward with time¹³, causing the locus of rock uplift and associated erosion to change. And the apparent match of modern contraction rates across the Himalaya¹⁴ with the measured slip over the past 9,000 years at the southernmost edge of

the Himalaya¹⁵ does not require an active fault in the middle of the range.

Furthermore, the spatial pattern of erosion rates is complicated. Wobus and colleagues' cosmogenically determined rates are only high within a few kilometres of the proposed fault and decrease northwards to background rates. In the next valley to the west (the Marsyandi), however, cosmogenic erosion rates are 4–20 times higher again in the Greater Himalaya, north of the physiographic transition and the Main Central Thrust¹⁶. Although the rates of cooling and erosion inferred from ⁴⁰Ar/³⁹Ar ages by Wobus *et al.* indicate an abrupt increase in long-term erosion rates at the proposed fault, the ages tend to get still younger to the north, implying faster rates. In the Marsyandi valley, rates of bedrock cooling imply a 2–10-fold increase in long-term erosion rates in the Greater Himalaya⁵ compared with the ages reported from the Lesser Himalaya just south of the Main Central Thrust^{5,6}. Moreover, higher rates in the Greater Himalaya are sustained despite a fourfold decrease in monsoonal precipitation⁵, indicating that the coupling of climate erosion and tectonics may be weaker than theory might predict².

Nonetheless, the cosmogenic data define an abrupt jump in millennial-scale erosion rates at exactly the same place as the offset in ⁴⁰Ar/³⁹Ar ages¹⁷. Unless landslides occurred to perturb these short-term rates¹⁶, the spatial coincidence of jumps in erosion rates at very different timescales is strong evidence that there is indeed an active, surface-breaking fault in the core of the Himalayan orogen. The lateral extent of this fault, any causal relationship to climate gradients and its role in proposed flow in the lower crust remain to be explored. ■

Douglas W. Burbank is at the Institute for Crustal Studies, Department of Geological Sciences, University of California, Santa Barbara, California 93106, USA.

e-mail: burbank@crustal.ucsb.edu

1. Beaumont, C., Fullsack, P. & Hamilton, J. in *Thrust Tectonics* (ed. McClay, K. R.) 1–18 (Chapman & Hall, London, 1992).
2. Beaumont, C., Jamieson, R. A., Nguyen, M. H. & Lee, B. *Nature* **414**, 738–742 (2001).
3. Willett, S. D. *J. Geophys. Res.* **104**, 28957–28982 (1999).
4. Dadson, S. J. *et al.* *Nature* **426**, 648–651 (2003).
5. Burbank, D. W. *et al.* *Nature* **426**, 652–655 (2003).
6. Wobus, C., Heimsath, A., Whipple, K. & Hodges, K. *Nature* **434**, 1008–1011 (2005).
7. Armstrong, F. C. & Oriol, S. S. *Am. Assoc. Petrol. Geol. Bull.* **49**, 1847–1866 (1965).
8. Harrison, T. M. *et al.* *Earth Planet. Sci. Lett.* **146**, E1–E7 (1997).
9. Copeland, P. *et al.* *J. Geophys. Res.* **96**, 8475–8500 (1991).
10. Dahlen, F. A. *Annu. Rev. Earth Planet. Sci.* **18**, 55–99 (1990).
11. Nelson, K. D. *et al.* *Science* **274**, 1684–1688 (1996).
12. Hodges, K. V. *Geol. Soc. Am. Bull.* **112**, 324–350 (2000).
13. Bollinger, L. *et al.* *Tectonics* **23**, TC5015 doi:10.1029/2003TC001564 (2004).
14. Wang, Q. *et al.* *Science* **294**, 574–577 (2001).
15. Lave, J. & Avouac, J. P. *J. Geophys. Res.* **105**, 5735–5770 (2000).
16. Niemi, N. A. *et al.* *Earth Planet. Sci. Lett.* (submitted).
17. Wobus, C. W., Hodges, K. V. & Whipple, K. X. *Geology* **31**, 861–864 (2003).

Metabolism

A higher power for insulin

Fiona M. Gribble

Glucose output from the liver is tightly regulated by insulin. But insulin holds sway over more than the liver — an unappreciated circuit in glucose control involves the opening of ion channels in the brain.

Developed countries are currently witnessing a surge in the incidence of adult-onset (type 2) diabetes, driven by soaring levels of obesity. Insulin, the hormone that normally coordinates the disposal of glucose after a meal, becomes less effective in obesity, resulting in the high blood-glucose concentration that is the hallmark of diabetes. Insulin is known to target the liver directly, but Pocai *et al.*¹ (page 1026 of this issue) identify an additional pathway by which the hormone, acting on the brain, sends messages via the nervous system to exert a higher level of control on glucose output from the liver.

The liver has a central role in glucose homeostasis because it extracts glucose from the bloodstream in times of plenty, and synthesizes glucose in times of need. Thus, it buffers the body from extremes in glucose concentration — the relative excess after meals, and the relative shortage between meals, particularly overnight and during periods of fasting. The liver recognizes these different energy states through changes in the blood insulin concentration; insulin binds to its receptor in the liver and directly inhibits glucose production. But this well-established pathway is far from the whole story.

In recent years, the idea that glucose metabolism throughout the body is coordinated by the brain has gained growing support². It is increasingly evident, for example, that the liver receives control signals from the hypothalamus, an area of the brain known to detect and integrate metabolic signals. Now, Pocai *et al.*¹ intriguingly show that insulin, in addition to its direct effects on the liver, also acts on specialized ion channels called K_{ATP} channels in the hypothalamus to control glucose production (Fig. 1). These channels, which lie in the outer membranes of hypothalamic neurons, were named for their ability to release potassium ions from cells in response to a drop in ATP (a molecule that provides energy for many cellular reactions and which is produced as glucose is metabolized). The authors' results link two findings that had previously been viewed in isolation: that insulin can modulate liver glucose output through an unknown action in the hypothalamus³, and that it opens K_{ATP} channels in an unidentified set of neurons in a part of the hypothalamus that monitors blood metabolites and hormones⁴.

K_{ATP} channels are generally known for

their ability to convert metabolic signals into changes in electrical activity in excitable cells such as neurons, cardiac muscle cells and some endocrine cells — including the insulin-producing β -cells of the pancreas. The channels close in response to ATP and sulphonylureas (a class of drugs used in the treatment of type 2 diabetes), and are opened by ADP (a metabolic product of ATP) and by certain lipids and insulin. K_{ATP} channels regulate both the resting membrane potential and cell excitability. Opening the channels — by, for example, decreasing the relative amount of ATP in response to falling glucose levels — allows potassium ions to leave the cell, resulting in membrane hyperpolarization and

reduced membrane excitability, which dampens its electrical activity.

It is known that insulin exploits a K_{ATP} -channel-dependent mechanism to electrically silence a subgroup of glucose-responsive neurons in the hypothalamus⁴. It seems contradictory that insulin, the hormonal indicator of nutrient excess, should have the same effect as glucose deprivation on K_{ATP} -channel activity and neuronal excitability. However, Pocai *et al.*¹ show convincingly in rats that opening K_{ATP} channels in the hypothalamus by infusing the K_{ATP} -channel opener diazoxide causes a neurally mediated reduction in liver glucose production, typical of the response to high rather than low nutrient availability. To substantiate the physiological relevance of this finding, they performed an exhaustive series of experiments on rats and mice. They infused insulin into the cerebral ventricles, hypothalamus or peripheral bloodstream, in the presence or absence of K_{ATP} -channel inhibitors in the brain, or in animals lacking subunits of the K_{ATP} channel, and arrive at a simple conclusion: a rise in physiological insulin concentrations can switch off liver

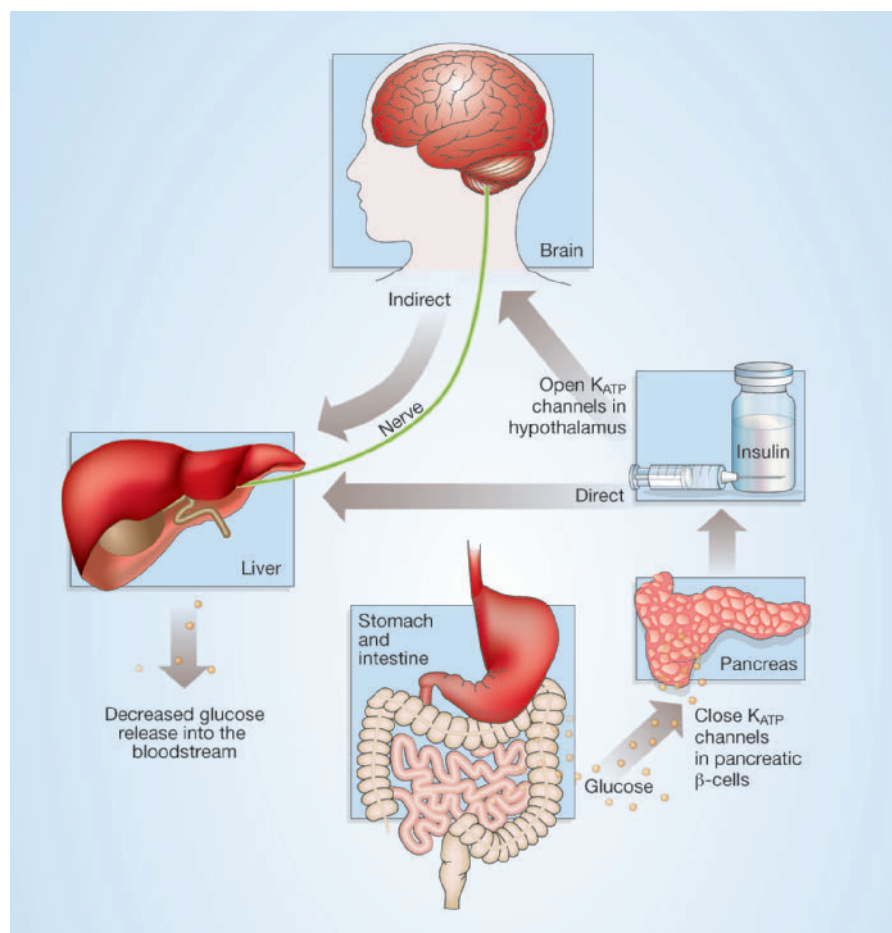


Figure 1 Control of liver glucose production after a meal. Glucose absorbed from the intestine stimulates the release of insulin into the bloodstream, which in turn inhibits liver glucose production by both a classical direct pathway and a newly identified indirect pathway. The indirect pathway involves the opening of ATP-sensitive potassium ion channels (K_{ATP} channels) in the outer membranes of neurons in the hypothalamus, resulting in signal transmission to the liver through the vagus nerve. This contrasts with the concomitant inhibition of K_{ATP} channels in pancreatic β -cells that underlies the glucose-stimulated release of insulin.



100 YEARS AGO

"The treatment of cancer with radium"
— The discovery of radium was speedily followed by its use in the treatment of cancer, and it was hoped that at last a remedy had been found for this terrible disease... A very large number of cases of cancer have been treated by radium in this country, on the Continent, and in America. Some have improved remarkably, but in most instances there has been no apparent benefit, and in no case has sufficient time elapsed to speak with certainty of cure... Radium is applied in small tubes to the surface of a tumour, and in some cases it has been found possible to place it in the interior of a growth through a small incision. The quantities available are so minute that only a small area can be treated at one time... Fortunately the radium can be used again and again, for its energy appears practically to be inexhaustible.
From *Nature* 20 April 1905.

50 YEARS AGO

"Biological hazards of radiation" — In the same week as the "Statement on Defence, 1955" and the White Paper "A Programme for Nuclear Power" were published in Great Britain, the Atomic Energy Commission of the United States produced a long and detailed report on the radiation and 'fall-out' produced by thermonuclear explosions. In the following week, the Federation of American Scientists communicated proposals for an international study of the potential dangers in tests of nuclear weapons... When on March 10 the Prime Minister was asked in the House of Commons to take the initiative and propose the establishment of a United Nations Commission... Sir Winston replied that experts at the disposal of the Government saw no reason to dissent from the opinion of the United States Atomic Energy Commission's report, issued on February 15, that the tests so far carried out have only released enough radioactive material into the atmosphere to cause an individual living in the United States to receive the same quantity of radiation that he would have received in an X-ray chest examination at a hospital... in making its proposal for an immediate study of the potential genetic risks of nuclear weapons, the Federation of American Scientists appeared to be envisaging dangers from a number of tests greatly in excess of those likely to be carried out in the foreseeable future.
From *Nature* 23 April 1955.

glucose production by opening hypothalamic K_{ATP} channels.

The implications of these findings deserve reflection, although the contribution of this pathway to the control of hepatic glucose output in intact animals, and in humans in particular, remains to be established. There are several inherited insulin-secretion disorders that are caused by altered K_{ATP} -channel activity in pancreatic β -cells, and assessment of liver glucose handling in such disorders would be of great interest. The authors raise the possibility that hypothalamic resistance to insulin might contribute to the raised liver glucose output typical of type 2 diabetes, suggesting that agents targeting such resistance (or even targeting hypothalamic K_{ATP} channels themselves) might be useful in treating this

condition. On a more sobering note, however, K_{ATP} -channel inhibitors are used routinely to increase insulin secretion in type 2 diabetes, and it may be that, if these oral hypoglycaemic agents gain access to and inhibit hypothalamic K_{ATP} channels, they might adversely affect glucose handling in the liver. There are clearly a number of critical studies needed to address these issues. ■

Fiona M. Gribble is at the Cambridge Institute for Medical Research and the Department of Clinical Biochemistry, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2XY, UK.

e-mail: fmg23@cam.ac.uk

1. Pocai, A. *et al.* *Nature* **434**, 1026–1031 (2005).
2. Schwartz, M. W. & Porte, D. Jr *Science* **307**, 375–379 (2005).
3. Obici, S., Zhang, B. B., Karkanias, G. & Rossetti, L. *Nature Med.* **8**, 1376–1382 (2002).
4. Spanswick, D., Smith, M. A., Mirshamsi, S., Routh, V. H. & Ashford, M. L. *Nature Neurosci.* **3**, 757–758 (2000).

Techniques

NMR on a chip

Robert Tycko

If a nanoscale gallium arsenide structure is excited with an oscillating magnetic field, superpositions of nuclear spin states can be created and detected electrically. Quantum computing could be the beneficiary.

Nuclear magnetic resonance (NMR) spectroscopy is a near-ubiquitous technique in the physical and biological sciences. Applications include the verification of molecular structures in synthetic chemistry, investigation of superconductivity in solid-state physics, the determination of protein conformations in structural biology, and diagnostic imaging in medicine. Typically, an adequate signal-to-noise ratio in NMR can be achieved only with a relatively large sample of 10^{15} – 10^{18} molecules or atoms. On page 1001 of this issue, Yusa *et al.*¹ demonstrate an unconventional approach to NMR in a cleverly designed semiconductor nanostructure with which strong signals can be detected from around 10^8 atoms.

NMR depends on manipulations of the spin states of atomic nuclei that possess magnetic moments. When an external magnetic field is applied to such nuclei, the energy levels allowed by the rules of quantum mechanics split — in the simplest case of a nucleus with a spin of magnitude 1/2, they split into two levels close to one another. If resonant radio-frequency pulse sequences with precisely controlled amplitudes, phases, frequencies and timing are then applied, transitions can be induced between these energy levels, creating coherent superpositions of spin quantum states. The frequencies at which these transitions occur ('NMR spectra') correspond to the gaps between spin energy levels and are therefore characteristic of a particular nucleus and its environment.

It is the ability to control quantum mechanical states more precisely than in any other type of spectroscopy that makes NMR spectroscopy useful in many disparate disciplines. The 'controllability' of nuclear spin states has also made NMR of interest in quantum computation (for an example see ref. 2). Here, mathematical problems can be efficiently solved by algorithms that depend on the manipulation of coherent superpositions of quantum mechanical states.

The main weakness of conventional NMR, which uses metal coils with centimetre-scale dimensions to detect nuclear-spin magnetism from samples with similar dimensions, is its relatively low sensitivity. Not only have Yusa *et al.*¹ significantly reduced the minimum sample needed to perform NMR spectroscopy, but they have also shown that, by exciting the nanostructure with radio-frequency magnetic fields, arbitrary superpositions of spin states can be created. These include 'multiple quantum coherences', superpositions of states that differ by more than one quantum of angular momentum³ in nuclei with a total spin greater than a half. Although this work is motivated by the goal of developing a solid-state electronic device for quantum computation, other applications and extensions of this nanotechnological approach to NMR may be possible.

A schematic representation of the device described by Yusa *et al.* is shown in Figure 1. Using a combination of semiconductor-growth, lithography and electrical-gating

techniques, they created a nanometre-scale channel of electron-doped gallium arsenide through which a current I of around 7 nanoamperes passes in response to an applied voltage V . At a temperature of about 100 millikelvin and with a static magnetic field B_0 of 5.5 tesla, the current generates large polarizations in the nuclear spin—that is, a large difference between the number of gallium and arsenic nuclei that have spin parallel to the static field and the number of nuclei with spin antiparallel to the field. As shown in earlier studies^{4–7}, under these experimental conditions the polarized nuclear spins interact with flowing electrons in such a way that the electrical resistance $R = V/I$ of the gallium arsenide channel increases by roughly 10%.

Simulations performed by Yusa *et al.* to fit their data indicate that this resistance change is directly proportional to the nuclear-spin polarization¹. When an oscillating magnetic field B_1 is applied with a frequency f_1 close to an NMR frequency, the direction of the nuclear-spin polarization is inverted in a periodic, sinusoidal manner, producing sinusoidal variations in R . By 'sweeping' the frequency of the applied field, the complete NMR spectrum of the gallium arsenide channel can thus be mapped out—not just the single-quantum transitions of the ⁶⁹Ga and ⁷⁵As nuclei seen in conventional NMR spectra, but also transitions involving states of these nuclei with spin 3/2 that differ by two or three quanta of angular momentum. Moreover, the sinusoidal variations in R persist over timescales approaching 1 millisecond, showing that

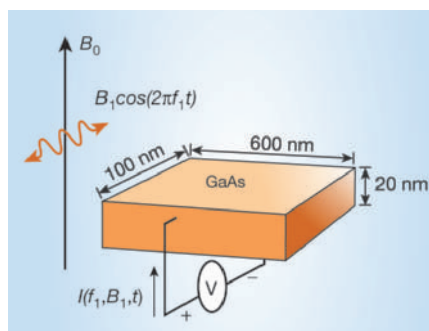


Figure 1 Nanoscale resonator. Current I passing through a gallium arsenide nanostructure as described by Yusa *et al.*¹ produces a large polarization in nuclear spin in the presence of a static magnetic field B_0 . Applying a second magnetic field B_1 , which varies sinusoidally with a frequency f_1 close to an NMR frequency, causes the spin polarization, and thus the resistance, in the GaAs channel to oscillate. Varying the frequency of this field allows the spectrum of all allowed NMR transitions to be mapped out electrically.

long-lived coherent superpositions are generated in these experiments, and that the creation of arbitrary superpositions would be possible with more elaborate radio-frequency irradiation schemes.

Could electrical detection using similar devices become a general method for performing NMR measurements? The electronic properties of gallium arsenide, and the technology to use these properties for creating complex nanostructures, are rather special. But nanostructures based on gallium arsenide are widely used for studying the

generic characteristics of confined, interacting electrons⁸. Nanometre-scale NMR devices such as those described by Yusa *et al.* could therefore contribute to our understanding of such systems and of interacting many-body systems in general. Given the interest in the development of organic materials possessing electronic properties analogous to those of gallium arsenide, it also seems likely that we could soon see nanometre-scale NMR devices constructed from organic molecules. By depositing biological macromolecules on such a device, and taking advantage of NMR tricks that permit the exchange of nuclear-spin polarization between molecules in contact with one another⁹, it might be possible to record NMR spectra of picomolar quantities of proteins and nucleic acids by purely electrical measurements.

Robert Tycko is in the Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, USA.

e-mail: robertty@mail.nih.gov

1. Yusa, G., Muraki, K., Takashina, K., Hashimoto, K. & Hirayama, Y. *Nature* **434**, 1001–1005 (2005).
2. Cappellaro, P. *et al. Phys. Rev. Lett.* **94**, 020502 (2005).
3. Vega, S. & Naor, Y. *J. Chem. Phys.* **75**, 75–86 (1981).
4. Kane, B. E., Pfeiffer, L. N. & West, K. W. *Phys. Rev. B* **46**, 7264–7267 (1992).
5. Wald, K. R., Kouwenhoven, L. P., McEuen, P. L., Vandervort, N. C. & Foxon, C. T. *Phys. Rev. Lett.* **73**, 1011–1014 (1994).
6. Machida, T., Yamazaki, T., Ikushima, K. & Komiyama, S. *Appl. Phys. Lett.* **82**, 409–411 (2003).
7. Yusa, G., Hashimoto, K., Muraki, K., Saku, T. & Hirayama, Y. *Phys. Rev. B* **69**, 161302 (2004).
8. Tycko, R., Barrett, S. E., Dabbagh, G., Pfeiffer, L. N. & West, K. W. *Science* **268**, 1460–1463 (1995).
9. Pines, A., Gibby, M. G. & Waugh, J. S. *J. Chem. Phys.* **59**, 569–590 (1973).

Plant biology

Scented story

Poets are captivated by the fragrance of flowers. So are scientists. A rather prosaic answer to the source of fragrance—for one species at least—is that it has a lot to do with the *ODORANT1* gene (see *Plant Cell* doi:10.1105/tpc.104.028837; 2005).

Julian C. Verdonk *et al.* have studied *Petunia hybrida* (pictured), which emits its scent at night to attract its nocturnal pollinators, hawkmoths. The scent consists of a combination of volatile benzenoid compounds. The chain of precursors from which benzenoids are produced is well understood, but what's less clear is why their synthesis and release peak at night.

To find out, Verdonk *et al.* compared the genes expressed in flowers that were producing scent with those from flowers that were just about to produce scent and

those from a different *P. hybrida* cultivar that emits low levels of benzenoids. This led the authors to a gene that they have called *ODORANT1* (*ODO1*), whose activity is increased in the fragrant flowers at the appropriate time of day.

Sequence comparisons suggested that the protein encoded by *ODO1* is a gene regulator. Indeed, eliminating *ODO1* expression led to decreased expression of several other genes, whose protein products catalyse early steps in the benzenoid-synthesis pathway. Benzenoid production was also decreased. Curiously, however, there was no effect on flower colour, even though one of the benzenoid precursors is also involved in making floral pigments. That tallies with the idea that pigment production occurs at an earlier



developmental stage than scent synthesis, and so is unaffected by the later experimental interference with *ODO1* expression.

That won't be the end of this scented story, however. One obvious question is how *ODO1* itself is

turned on at the right time. Another is how its protein product interacts with its target genes—does it do so directly or through other factors? And can the scent of other plant species be attributed to *ODO1* activity?

Amanda Tromans

Neurobiology

Sculpted by competition

Ole Petter Ottersen

Neuronal competition helps connections to form in the brain: the branches of less active neurons are more likely to retract — and, it now seems, less likely to grow — than those of their more active neighbours.

Competition pervades everyday life and sculpts our society. It is also an essential factor in the sculpting of our brains. On page 1022 of this issue, Hua and colleagues¹ provide new insight into the rules that underlie this competition.

Our brains contain billions of nerve cells that are wired together in a specific manner. During brain development, these nerve cells extend thin processes — axons — that branch and establish contacts with other brain cells. The contacts, called synapses, are sites at which neurons communicate through the release of chemical neurotransmitters. One of the most fundamental but also most complex questions in neuroscience is how the appropriate connections are formed when the brain develops. A detailed map of the vast number of synapses in the adult brain cannot be laid out in our genes: our genome is simply too small to specify such a complicated network.

It is now well established that specific molecular cues act as signposts that guide growing axons to their target cells. But the refinement of the connections depends on neural activity. A wealth of data suggests that electrical activity in developing networks influences the branching of the terminal parts (the 'arbors') of axons, and hence the formation of functional neural circuits².

Current dogma, inspired by elegant studies of the development of peripheral motor axons, holds that activity moulds the axonal arbor by regulating the extent and rate of branch retractions. Hua *et al.*¹ add to this by providing evidence that activity can also control the extent and rate of branch formation. Furthermore, their findings point to a new set of rules for activity-based competition between neighbouring axons: less active axons lose out in the competition with their more active peers by being denied the possibility of fully developing their terminal arbors.

To look at the effects of neural activity on branch formation, Hua *et al.* used the retino-tectal pathway in zebrafish as a model system (Fig. 1). This pathway is involved in the control of eye movements, and is formed by the axons of retinal ganglion cells. By transfecting individual ganglion cells with a fluorescent marker (whereby genes expressing the marker are transferred to the cell using a virus), the authors could monitor the growth and branching of single retinotectal axons by two-photon microscopy in live zebrafish. Two-photon microscopy is well suited

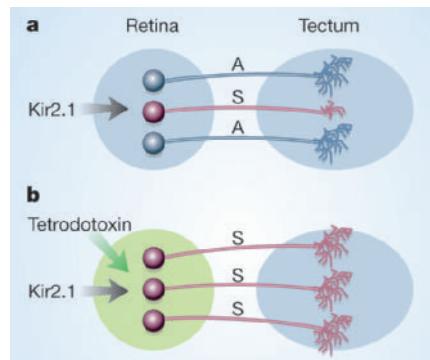


Figure 1 Competition in the brain. Retinal ganglion cells convey information from the retina to the optic tectum in the zebrafish brain. **a**, Individual retinal ganglion cells can be silenced by transfection with the potassium channel Kir2.1. Hua *et al.*¹ have found that the silenced axon (S) grows a less elaborate terminal arbor than do adjacent, active axons (A). **b**, Silencing of all retinal ganglion cells by injecting the neurotoxin tetrodotoxin into the eye restores the ability of the Kir2.1-expressing cell to grow a terminal arbor. Thus, competition between neighbouring axons plays a part in the growth and branching of axon arbors.

for such analyses^{3,4} because it allows many images to be obtained without interfering unduly with the growth process.

To investigate how inactive axons fared in the competition with active neighbouring ones, Hua *et al.* needed to suppress the activity of single axons. To this end, they took advantage of previous findings⁵ showing that individual neurons can be silenced by transfecting them with a specific type of potassium channel — the human inward rectifier K⁺ channel Kir2.1. Transfected cells are silenced because overexpression of Kir2.1 causes them to hyperpolarize. Hua and colleagues' key observation was as follows: axons of cells that were made to overexpress Kir2.1, and which were presumably inactive, showed reduced growth and produced less elaborate terminal arbors compared with the axons of adjacent cells (Fig. 1a).

In itself, this finding cannot be taken as evidence of competition: growth could have been impeded as a direct consequence of inactivation. This possibility was ruled out by a crucial experiment in which all axons of the retinotectal pathway were silenced by tetrodotoxin — a neurotoxin isolated from puffer fish. It turned out that this neurotoxin restored the ability of the Kir2.1-expressing

axons to grow terminal arbors (Fig. 1b). So the growth inhibition of individually silenced Kir2.1-expressing axons must be due to negative feedback — directly or indirectly — from nearby, active axons.

What could be the mechanism that explains why Kir2.1-expressing neurons lose this growth competition with normal cells? An obvious possibility is that the effects of Kir2.1 overexpression are mediated through neuronal hyperpolarization, and a consequent block of neurotransmitter release. In support of this, Hua *et al.* found that the inhibitory effect of Kir2.1 on axon growth could be reproduced by transfecting retinal ganglion cells with a modified synaptic protein that suppresses the recycling of synaptic vesicles (a crucial stage in synaptic neurotransmission). Later-acting mechanisms are less clear. One possibility is that inactive axons fall prey to inhibition by adjacent, more active axons through the release of retrograde signals from the axons' target cells in the tectum (the dorsal part of the midbrain).

Hua and colleagues' main conclusion is that activity-based competition between axons sets in at an early stage, before the axons have developed a full terminal arbor. The axons seem forced to obey the rule that if you are less active than your neighbour, you are not allowed to grow.

These findings do not overthrow the current dogma — that competition acts by way of axon retraction or elimination. Rather, they suggest that the picture is more complicated, and that relative activity influences the development of functional networks by affecting axon growth as well as axon regression. These conclusions must, however, be tempered by the fact that the experimental design — based on silencing individual axons — does not mirror the situation in the developing brain, where neighbouring axons are likely to compete on more equal terms.

Nonetheless, through their ingenious combination of techniques, Hua *et al.*¹ have provided fresh insight into the mechanisms that sculpt the developing brain. One is left wondering to what extent this newly discovered rule for axon competition applies to other neuronal pathways and other species — and whether this rule is also relevant to the ability of the mature brain to develop and change.

Ole Petter Ottersen is at the Centre for Molecular Biology and Neuroscience, Institute of Basic Medical Sciences, Domus Medica (Sognsvannsveien 9), University of Oslo, N-0317 Oslo, Norway.
e-mail: o.p.ottersen@medisin.uio.no

- Hua, J. Y., Smear, M. C., Baler, H. & Smith, S. J. *Nature* **434**, 1022–1026 (2005).
- Katz, L. C. & Shatz, C. J. *Science* **274**, 1133–1138 (1996).
- Grutzendler, J., Kasthuri, N. & Gan, W.-B. *Nature* **420**, 812–816 (2002).
- Trachtenberg, J. T. *et al.* *Nature* **420**, 788–794 (2002).
- Burrone, J., O'Byrne, M. & Murthy, V. N. *Nature* **420**, 414–418 (2002).

Obituary

Hans A. Bethe (1906–2005)

Were we asked to characterize Hans Bethe with one word, we would choose strength — strength of intellect and strength of character. From the moment he reached the frontier of physics research, it was evident that he possessed exceptional intellectual strength and the personal resources to put his talents to outstanding use. The true extent of his moral strength would only emerge when his role in the birth of the nuclear age posed a host of ethical choices.

Bethe, who died on 6 March, was born in 1906 in Strasbourg, Alsace-Lorraine — then part of the German Empire — and went to school and university in Frankfurt am Main. In 1926, he joined Arnold Sommerfeld's famous school of theoretical physics in Munich just as Erwin Schrödinger's wave mechanics theory was appearing. Bethe absorbed the new theory quickly because, as he recalled, he was not burdened by knowing the recipes of the 'old' quantum theory.

By 1931, his rapidly growing publication list included three classic papers: the spectrum of an atom in a host lattice, one of the first applications of group theory to quantum mechanics; the quantitative description of how charged particles lose energy when traversing matter, important to the interpretation of nuclear data; and the exact spectrum and wavefunctions of a linear chain of interacting spins. All three papers solved difficult problems, and all were marked by formidable mathematics deployed with virtuosity. In the solution to the spin problem, Bethe introduced the now-famous 'Bethe ansatz', a tool used by armies of theorists to this day.

Bethe's amazing ability to analyse and synthesize a huge and still-growing body of intricate knowledge was first displayed in the 1933 edition of the *Handbuch der Physik* in two encyclopaedic articles: on one- and two-electron systems; and, with Sommerfeld, on the theory of metals.

But 1933 was also a personal watershed for Bethe. His parents were Protestants, but his mother had been Jewish at birth, and so he was dismissed by the Nazis from his university post. He moved to England, but, after two very productive years there, left Europe to join Cornell University in New

York, and never looked back. He was to stay at Cornell, despite many offers, for the next 70 years. In 1947, when asked to succeed Sommerfeld in Munich, he declined, writing that he felt "as if I was born in Germany only by mistake, and only came to my true homeland at age 28".

Bethe's involvement with nuclear

written in part with experimental colleagues at Cornell, and for years an indispensable text for students and experts alike.

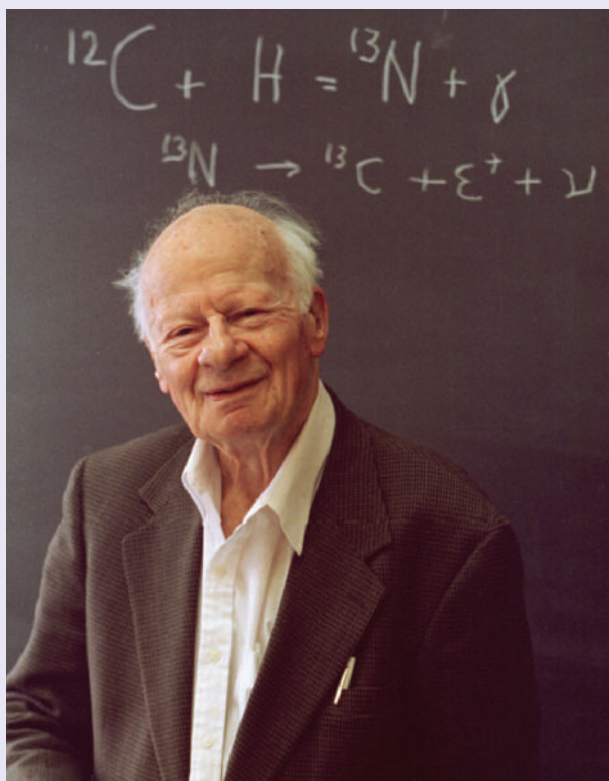
That nuclear processes must be responsible for energy production in stars was accepted by the 1930s, but specific reactions that could account for observations remained to be identified.

The first step, following a suggestion by Carl-Friedrich von Weizsäcker, was Bethe's 1938 paper with Charles Critchfield on the fusion of two protons into deuterium and the subsequent reactions that eventually lead to helium. Their results seemed to be in conflict with the solar models of the day; after the models were corrected, the agreement was close. But the proton reaction could not account for energy production by more massive stars. This led Bethe to look into other processes for converting hydrogen into helium, resulting in his discovery of the carbon–nitrogen cycle, in which carbon is not destroyed but is used repeatedly as a catalyst. His monumental carbon-cycle paper was published in 1939, and led in 1967 to Bethe receiving the first Nobel prize on a topic in astronomy.

The discovery of fission in 1938 and Germany's invasion of Poland ended what Bethe called "the happy thirties" of physics. Well before the Japanese attack on Pearl Harbor in 1941, Bethe and his good friend Edward Teller were eager to help with the defence of their new home. As they were not yet allowed

to do secret work, they asked Theodore von Kármán what they should do. The result — "Deviations from thermal equilibrium in shock waves", the analysis of the role of molecular structure in fluid dynamics under extreme conditions — was immediately stamped 'secret'. Bethe was very proud of this work, written in part in a mountain cabin during the Bethes' annual summer pilgrimage to the West, and only declassified long after the war.

Bethe participated in secret studies of both the fission and fusion weapon concepts before the start of the Manhattan Project to develop the atomic bomb. But, being a supremely practical man, he at first doubted that nuclear weapons could be invented in time to influence the course of



Giant of physics who grappled with the moral implications of nuclear weapons.

physics began in England in collaboration with Rudolf Peierls, his close friend since their days as students of Sommerfeld. At this time he devised the first theory of order–disorder phase transitions that took short-range correlations between atoms into account. Also in England, with Walter Heitler, he calculated the cross-section for electron–positron pair production, which played a central role in cosmic-ray physics and the discovery of mesons.

Bethe's mastery of the rapidly expanding field of nuclear physics produced innumerable invitations to speak, and he decided it would be more efficient to write a review than to spend so much time travelling across North America in trains. The result was the 'Bethe Bible', three book-length issues of *Reviews of Modern Physics*,

the war. Robert Oppenheimer and Teller changed his mind, and he went to the US government's newly founded Los Alamos nuclear laboratory in 1943 where he became head of the Theoretical Division. Because so many of the phenomena involved in nuclear explosions were not yet accessible to empirical study, theoretical physics was indispensable to the bomb project, especially after it turned out that a plutonium bomb would have to use a tricky implosion mechanism. Bethe, with his unequalled command of physics and serene but commanding persona, became the maestro of Oppenheimer's theoretical orchestra — which did not lack prima donnas.

At the end of the war, Bethe returned to Cornell, bringing Richard Feynman and Philip Morrison with him. Robert Wilson, who had led the experimental nuclear physics division at Los Alamos, and other veterans of the bomb project soon joined them. This established Cornell as a leading centre of theoretical and experimental particle physics, with Bethe as teacher and mentor of scores of students and postdocs. He himself started the spectacular postwar development of quantum electrodynamics (QED), the quantum-mechanical description of the electromagnetic interaction. This he did with the first, somewhat slapdash but essentially correct, calculation of radiative corrections to the hydrogen spectrum. He worked it out on the train ride from the conference where Willis Lamb's experiment disagreeing with QED, as the theory was then understood, was first announced. QED was then taken over by a brilliant young generation: Feynman and Freeman Dyson at Cornell, and Julian Schwinger at Harvard.

But there could be no return to the ivory tower of the "happy thirties". For so central a character in the drama that culminated in the bombing of Hiroshima and Nagasaki, politics was bound to be an inescapable presence. In this setting, Bethe assumed a variety of roles: as an expert technical consultant in further weapons development, and an opponent to such development; as an adviser to the highest levels of government, and an outspoken critic of government policy; and as a defender of colleagues under attack by influential demagogues. That he played all these roles shows how difficult he found it to square his deep moral and ethical convictions, and his desire to influence fateful events, with the realities of the

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Presidential adviser: Bethe receives the Enrico Fermi medal from President John F. Kennedy in 1961. Bethe's wife Rose, his companion through 66 years of marriage, looks on.

context in which he now lived. As he put it, "sometimes I wish I were a more consistent idealist".

Bethe never regretted his role in the wartime bomb project, for he had feared that Nazi Germany was developing such a weapon; but he was deeply worried by what he had helped to unleash. Soon after 1945, he joined Albert Einstein and others in a public-education campaign. He publicly opposed the development of a hydrogen bomb on moral grounds, but participated in its research in the hope that he could prove it impossible. After Stanislaw Ulam and Teller discovered how it could be done, he remained involved because he believed the Soviets would also invent it. The H-bomb controversy finally led to the sham process that removed Oppenheimer's security clearance, and to his irrevocable split with Teller.

As a member of the President's Science Advisory Committee (PSAC) from its start in 1957, Bethe advocated a ban on nuclear tests. After a long struggle, in which he played a central part, this came to partial fruition in 1963 with the atmospheric test ban. PSAC studies convinced Bethe that it was futile to attempt a defence against ballistic missiles. After President Lyndon Johnson rejected this assessment, he and Richard Garwin went public with their views. Bethe was to take the same vigorous public role in opposing President Ronald Reagan's 'Star Wars' plans. And in Los Alamos on the fiftieth anniversary of Hiroshima, he called on scientists everywhere to desist from work on creating and developing further nuclear weapons.

Between 1940 and 1970 Bethe did not work explicitly on astrophysics but had a strong indirect influence by supporting Willy Fowler, who was turning nuclear astrophysics into an empirical science at the California Institute of Technology in

Pasadena. Bethe was the role model for Fowler's enterprise, interacting constantly with him and sending him young postdocs — including one of us (E.E.S.). During the early 1970s, Bethe worked on the equation of state of high-density nuclear matter, which is needed to understand the dynamics of neutron stars. Of particular interest was how matter 'stiffened' above normal nuclear densities.

Soon after his official retirement, Bethe returned to astrophysics and did cutting-edge research on three different topics until well beyond the age of 90. In close collaboration with Gerald Brown, a major focus was work on the formation of type II

supernovae. This process starts with the collapse of a massive star and ends with its outer layers exploding, leaving a neutron star at its centre. Their calculations made use of Bethe's work on high-density equations of state. Unresolved controversies remain because of the many competing mechanisms for forcefully reversing the collapse, including neutrino transport, convection carrying energy outwards and exothermic nuclear reactions in intermediate layers. One of Bethe's greatest contributions, exploiting ingenious skills he first developed at Los Alamos, was showing how auxiliary calculations could be done simply and semi-analytically.

Bethe also made important contributions, in part with John Bahcall, to unravelling the solar neutrino problem — the fact that the flux of neutrinos coming from the Sun is considerably smaller than solar models predict — after it became clear that the theory of the neutrino had to be modified to take neutrino oscillations into account. Later on, Bethe and Brown studied various kinds of binary stars, including the merger of a neutron star and a small black hole, which is of interest for gravitational wave detectors.

The accomplishments of Hans Bethe were prodigious; it is impossible to do justice to them here. We offer this as a snapshot of him as a scientist devoted to the country that had embraced him, even though its current path disturbed him deeply, and as a man who loved his students and colleagues, his children and his wife Rose — his constant companion and foremost adviser in all the difficult decisions he had to face. **Kurt Gottfried and Edwin E. Salpeter**

*Kurt Gottfried and Edwin E. Salpeter are in the Department of Physics, Cornell University, Ithaca, New York 14853-5001, USA.
e-mail: kg13@cornell.edu*

Earth science

Heated exchange on the coast

Geology **33**, 245–248 (2005)

Volcanologists are continuing to dig into the consequences of eruptions of the Soufrière Hills volcano, Montserrat, in the Caribbean. The volcano began to erupt in 1995, but an especially severe episode occurred on 13 July 2003. Marie Edmonds and Richard A. Herd have examined deposits to the northeast of the volcano, and conclude that the devastation there was the result of the explosive combination of sea water and pyroclastic flows from the volcano.

Pyroclastic flows are among the deadliest results of a volcanic eruption. They are intensely hot mixtures of gas, ash and rocks, which travel downslope at great speed. On reaching the coast, the flow of 13 July continued underwater and generated a tsunami. But Edmonds and Herd believe that the interaction with sea water also produced a reverse, inflated pyroclastic surge, which flowed along the coast and inland. They estimate that

15 July 2003: the Soufrière Hills volcano still steams after the violent activity two days previously.

it caused destruction at least 25 m above the ground, reaching 4 km inland and 300 m above sea level.

Such 'coastal explosions' can be

simulated experimentally. But, given that coastlines are especially prone to erosion, they are comparatively rare in the geological record.

Tim Lincoln

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Pharmacology

Antidepressant action

Neuron **46**, 65–74 (2005)

Antidepressants such as Prozac are thought to work by interfering with the conversation between neurons that transmit the neurotransmitter serotonin and those that receive it. Fu-Ming Zhou *et al.* suggest that such drugs also have an effect on dopamine-producing cells.

Neurons generally use chemical neurotransmitters to talk to each other, the neurotransmitters then being taken back up into the transmitting cell. In the striatum of the mammalian brain — a region involved in processes including motor coordination and cognition — 'dopaminergic' nerve terminals, which release dopamine, are common, and serotonin-releasing terminals are more modestly represented. Generally, these nerve terminals take up only the particular neurotransmitter that they make.

But Zhou *et al.* find that increasing the serotonin levels in mouse striatal slices leads to serotonin uptake into dopaminergic cells. The authors further show that the dopaminergic cells then release both dopamine and serotonin. Moreover, applying fluoxetine (Prozac) — which inhibits serotonin uptake by serotonin-releasing cells — seems to have similar effects. If borne out, the results imply that such antidepressants may work in part by altering the balance of dopamine and serotonin release.

Amanda Tromans

Quantum dynamics

Reverse for simulation

Phys. Rev. Lett. **94**, 130401 (2005)

It may be possible to turn back time, at least in quantum-dynamical simulations of many-body systems. That's the conclusion of Mark Dowling and colleagues, who have studied the evolution of an ultracold Bose–Einstein condensate of atoms.

They used a fundamental property of the hamiltonian, the quantum-mechanical 'operator' that determines the energy states of a system. Evolution backwards in time under one hamiltonian is equivalent to evolution forwards under the same hamiltonian with the opposite sign. Changing the sign of the hamiltonian halfway through a simulation is therefore equivalent to reversing the direction of time.

Doing just this and seeing whether the initial conditions are regained is an effective check of the accuracy of a simulation, the authors suggest. Such checks are essential, but taxing — the complexity of many-body systems scales exponentially with the number of bodies involved, and small changes in initial conditions can produce large variations in outcome.

Dowling *et al.* tested their method on a system modelling a Bose–Einstein condensate with Avogadro's number (6.02×10^{23}) of atoms. The final state of their simulation accurately reproduced the initial state. The authors note that their condensate simulation could be

realized experimentally, and their method used to check similar quantum-dynamical calculations with applications not just in condensed-matter physics, but also in biology and astrophysics.

Richard Webb

Solid-state optics

Good vibrations

Appl. Phys. Lett. **86**, 154107 (2005)

To cool a block of material, you have to take heat out of it. There are various ways to do this. Shining laser light onto the material is one of the less intuitive methods — you might expect the material to heat up rather than cool down.

But certain solids respond to laser illumination by emitting light themselves. In most such luminescent systems, there is less power in the emitted light than in the incident light; in some, however — such as glasses doped with special absorbing ions similar to those used for optical fibres — the emitted light can gain energy from thermal vibrations in the solid, and so can be more energetic than the absorbed light. If non-radiative losses are kept low, the material, when illuminated, cools down.

J. Thiede *et al.* have now used this technique to cool a piece of glass fibre down to -65°C , comfortably beating the previous lowest temperature achieved for a solid using this method, -37°C . Optical refrigerators based on this technology could be of great practical importance owing to their compact dimensions and the absence of vibrations associated with moving parts, the authors say.

Andreas Tröbesinger

Arboreal ants build traps to capture prey

Tiny ants construct an elaborate ambush to immobilize and kill much larger insects.

To meet their need for nitrogen in the restricted foraging environment provided by their host plants, some arboreal ants deploy group ambush tactics in order to capture flying and jumping prey that might otherwise escape^{1–4}. Here we show that the ant *Allomerus decemarticulatus* uses hair from the host plant's stem, which it cuts and binds together with a purpose-grown fungal mycelium, to build a spongy 'galleried' platform for trapping much larger insects. Ants beneath the platform reach through the holes and immobilize the prey, which is then stretched, transported and carved up by a swarm of nestmates. To our knowledge, the collective creation of a trap as a predatory strategy has not been described before in ants.

Allomerus decemarticulatus (Myrmicinae) is specifically associated with the Amazonian ant-plant *Hirtella physophora* (Chrysobalanaceae), which houses colonies in leaf pouches. Workers build galleried structures on their host plant's stems in which they pierce numerous holes that are slightly larger in diameter than their heads, allowing them to enter and exit⁵ (Fig. 1a). First, they cut plant hairs (trichomes) along the stems, clearing a path. Then, using uncut trichomes as pillars, they build the gallery's vault by binding cut trichomes together with a compound that they regurgitate (for details, see supplementary information). Later, this structure is reinforced by the mycelium of a complex of sooty-mould species that has been manipulated by the ants. Fungal growth starts around the holes and then spreads rapidly to the rest of the structure.

We noted that the stems of 34 young seedlings, which had not yet developed leaf pouches, did not bear fungus; nine saplings raised in a greenhouse in the absence of *Allomerus* developed leaf pouches but never bore fungus. However, 15 saplings raised in the presence of ants bore mycelia, whose development was limited to the galleries. When we eliminated the associated ants from five of the 15, the fungus on the galleries grew into a disorganized structure, and none of the nine new stems that developed bore any fungus at all.

Because prey seemed to be immobilized on the top surface of these galleries, we investigated whether these structures could be acting as a trap. Our observations revealed that *Allomerus* workers hide in the galleries with their heads just under the holes, mandibles wide open, seemingly waiting for an insect to land. To kill the insect, they grasp its free legs, antennae or wings, and move in and out of holes in opposite directions until

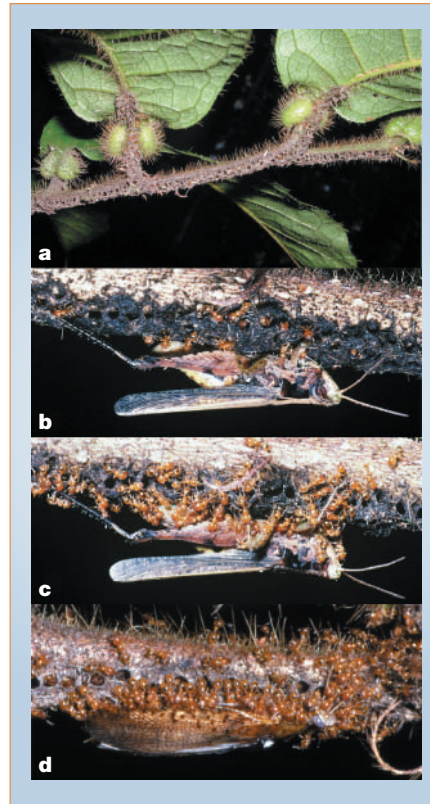


Figure 1 Galleried platforms built by *Allomerus* workers act as traps for prey. **a**, Base of the lamina of two *Hirtella physophora* leaves: the pouches serve as ant nests, which are interconnected by a gallery pierced with numerous holes. **b**, Start of the capture of a locust. **c**, One hour later, recruited workers have left the gallery in order to bite, sting and stretch the prey. **d**, Completion of a cricket capture.

the prey is progressively stretched against the gallery and swarms of workers can sting it (Fig. 1b–d). The ants then slide the prey over the top of the gallery — again moving in and out of holes, but this time in the same direction. They move it slowly towards a leaf pouch, where they carve it up.

Because the requirement for protein is crucial, some arboreal ants consume parts of the Hemiptera that they attend⁶, others rely on microsymbionts to recycle nitrogen⁷, and some have a thin cuticle and non-proteinaceous venom that economizes on nitrogen use⁸. Nevertheless, most of them must capture prey that land on their host plant⁵. Given this constraint, the tiny *A. decemarticulatus* workers have developed a tripartite association with their host plant and a fungus collectively to ambush their prey.

Alain Dejean*, Pascal Jean Solano*, Julien Ayroles†, Bruno Corbara‡, Jérôme Orivel*

*Laboratoire d'Évolution et Diversité Biologique (UMR CNRS 5174), Université Toulouse III,

31062 Toulouse, France

e-mail: orivel@cict.fr

†Department of Animal Biology, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801, USA

‡Laboratoire de Psychologie Sociale de la Cognition (ESA CNRS 6024), Université Blaise Pascal, 63037 Clermont-Ferrand, France

1. Heil, M. & McKey, D. *Annu. Rev. Ecol. Syst.* **34**, 425–553 (2003).
2. Hölldobler, B. & Wilson, E. O. *The Ants* (Harvard Univ. Press, Cambridge, Massachusetts, 1990).
3. Dejean, A. & Corbara, B. in *Arthropods of Tropical Forests. Spatio-temporal Dynamics and Resource Use in the Canopy* (eds Basset, Y., Novotny, V., Miller, S. E. & Kitching, R. L.) 341–347 (Cambridge Univ. Press, Cambridge, UK, 2003).
4. Morais, H. C. *Insectes Soc.* **41**, 339–342 (1994).
5. Benson, W. W. in *Amazonia* (eds Prance, G. & Lovejoy, T. E.) 239–266 (Pergamon, New York, 1985).
6. Offenberger, J. *Behav. Ecol. Sociobiol.* **49**, 304–310 (2001).
7. van Borm, S., Buschinger, A., Boomsma, J. J. & Billen, J. *Proc. R. Soc. Lond. B* **269**, 2023–2027 (2002).
8. Davidson, D. W. *Biol. J. Linn. Soc.* **61**, 153–181 (1997).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

Predation

Prey plumage adaptation against falcon attack

Several plumage types are found in feral pigeons (*Columba livia*), but one type imparts a clear survival advantage during attacks by the swiftest of all predators — the peregrine falcon (*Falco peregrinus*)^{1,2}. Here we use quantitative field observations and experiments to demonstrate both the selective nature of the falcon's choice of prey and the effect of plumage coloration on the survival of feral pigeons. This plumage colour is an independently heritable trait³ that is likely to be an antipredator adaptation against high-speed attacks in open air space.

The polymorphic feral pigeon is ideal for investigating the effects of different plumage combinations on predator success. One plumage phenotype — known as the 'wild' variant — is blue-grey but has a white rump between the base of the tail and the lower back, closely resembling the feral pigeon's rock-dove ancestor^{3,4} (Fig. 1); all other plumage types lack this contrasting rump patch³. We tested whether the wild coloration might impart a selective advantage over other plumage types during high-speed diving attacks by falcons, which often exceed 157 m s^{−1} (ref. 2; Fig. 1a).

During a seven-year period, we recorded 1,485 attacks by five adult peregrine falcons on free-ranging flocks of feral pigeons in Davis, California, during their daily linear

commute, and an additional 309 attacks when three of these falcons were juveniles. For each attack, we noted the plumage phenotype of the selected pigeon and whether the stoop led to a successful capture (for methods, see supplementary information).

To determine the natural distribution of plumage phenotypes, we banded and recorded data from 5,235 pigeons in the study area. Adult falcons selected plumage phenotypes for attack in the same relative proportions that occurred naturally (Fig. 1a, inset), except for the wild phenotype ($P=0.006$). Juvenile falcons, however, attacked all plumage phenotypes in proportion to their frequency in the population.

Comparison of adult and juvenile falcons revealed a significant difference in the plumage phenotypes they selected as prey ($F[5,1782]=2.129$, $P=0.05$; Fig. 1b). But there was no difference between adults and juveniles in their captured pigeons' pattern of plumage ($F[5,585]=0.39$, $P=0.85$; Fig. 1b), suggesting that both age classes learn to avoid the wild plumage during attacks.

Adults were significantly more adept hunters, capturing their prey on 40% of stoops, whereas juvenile success was only 19% ($P=0.0001$). Both adults and juveniles, however, showed a significant interaction between the phenotypes selected for attack and those captured (adult: $F[5,2075]=233.7$, $P=0.0001$; juvenile: $F[5,360]=2.66$, $P=0.02$; Fig. 1b). Much of this effect was driven by the wild phenotype, which accounted for only 2% of all captures in both age classes.

To confirm the advantage afforded by a white rump to pigeons during a high-speed attack, we captured 756 wild and blue-barred pigeons and switched their rump-patch feathers. We then released these pigeons and monitored predation by three adult peregrine falcons. After plumage transfer, we found that capture rates were reversed (Fig. 1c). The original wild phenotype now suffered predation at the same rate as the unmanipulated blue-barred type, whereas the manipulated blue-barred type now had rates of predation as low as the unmanipulated wild type ($P<0.0001$). This indicates that the dorsal white rump patch is important for the survival of feral pigeons during attacks by peregrine falcons.

All feral pigeons perform the same evasive roll during predation by falcons (Fig. 1a). The protective white patch may disguise the initiation of the pigeon's evasive roll by contrasting conspicuous (white patch) and cryptic targets (grey wings and body)⁵. A fast-flying falcon primed to a conspicuous target centered on the roll might fail to detect the dodge initiated by the cryptic wings as the predator closes from behind. When pursued by predators, schools of fish in open water and many shorebirds also display their dark dorsal and light ventral surfaces in a display that alternates between a cryptic and a

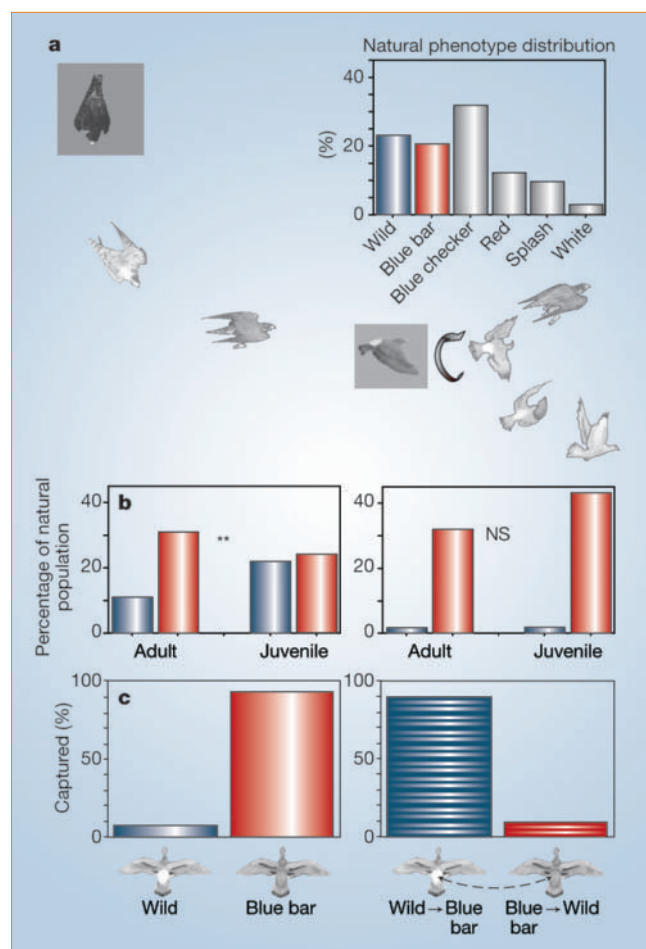


Figure 1 Adaptive coloration in pigeons hunted by peregrine falcons. **a**, Sequence of attack by a falcon as it closes on a pigeon: from the bottom of its dive (top left), the falcon arcs behind the pigeon to begin its 20–100-m horizontal approach; the pigeon then takes avoiding action by dipping one wing, rapidly rotating (arrow), and rolling out of the path of the falcon. Inset, natural distribution of the six pigeon plumage phenotypes in the study population ($n=5,235$). **b**, Percentage of wild-coloured (shown in blue) and blue-barred (shown in red) pigeons selected for attack (left panel) and actually captured (right panel) from the natural population by adult and juvenile falcons. (Statistics indicate comparison of all phenotypes.) **c**, Percentage of wild-coloured (white rump) and blue-barred pigeons captured in the natural population (left panel; $n=203$) and after plumage transfer (right panel; $n=69$) in a field experiment.

conspicuous signal^{6–9}. The use of contrasting patterns as an antipredator mechanism is widespread and may represent a case of convergent evolution.

Given such a selective advantage, why is the wild phenotype not better represented in flocks of feral pigeons? Unlike their monomorphic ancestor, the rock dove, feral pigeons often choose mates with plumage unlike their own (negative assortative mating), and this maintains plumage polymorphisms³. Also, the conspicuous white rump patch could prove disadvantageous when the bird is near cover, such as trees, or confronted by other predators, particularly accipiters and falcons that typically chase down their prey at lower speeds.

In the eastern part of the pigeon's native range in Eurasia, *C. livia* is sympatric with a subgenus of falcons, *Hierfalco*, that typically chase their prey in level flight. Pigeons from this region lack the white rump patch^{1,4}. The decline in recent decades of peregrine-falcon populations worldwide relaxed their selective pressure on feral pigeons, but this has now reversed, with falcons recolonizing many of their former haunts¹⁰. Over the study period, the proportion of wild plumage types in our study population increased significantly relative to blue-barred types ($P=0.01$), in parallel with a steady increase in predation by peregrine falcons. This suggests that, although

pigeon plumage polymorphism persists, falcon predation pressure can lead to an increase in the relative proportion of the wild pigeon phenotype.

Alberto Palleroni*‡, Cory T. Miller†, Marc Hauser*, Peter Marler‡

*Primate Cognitive Neuroscience Laboratory, Department of Psychology, Harvard University, Cambridge, Massachusetts 02138, USA
e-mail: aliparti@wjh.harvard.edu

†Laboratory of Auditory Neurophysiology, Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

‡Laboratory of Animal Communication, University of California, Davis, California 95616, USA

- Cade, T. J. *The Falcons of the World* (Cornell Univ. Press, Ithaca, New York, 1982).
- Tucker, V. A. *J. Exp. Biol.* **201**, 403–414 (1998).
- Johnston, R. F. & Janiga, M. *Feral Pigeons* (Oxford Univ. Press, Oxford, UK, 1995).
- Goodwin, D. *Pigeons and Doves of the World* (Cornell Univ. Press, Ithaca, New York, 1970).
- Bond, A. B. & Kamil, A. C. *Anim. Learn. Behav.* **27**, 461–471 (1999).
- Helfman, G., Collette, B. & Facey, D. *The Diversity of Fishes* (Blackwell Science, Malden, Massachusetts, 1997).
- Tinbergen, N. *The Study of Instinct* (Oxford Univ. Press, London, 1951).
- Brooke, M. de L. *Funct. Ecol.* **12**, 339–346 (1998).
- Lima, S. L. *Wilson Bull.* **105**, 1–47 (1993).
- Cade, T., Enderson, J. H., Thelander, C. G. & White, C. M. (eds) *Peregrine Falcon Populations: Their Management and Recovery* (The Peregrine Fund, Boise, Idaho, 1988).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.

Centennial-scale climate cooling with a sudden cold event around 8,200 years ago

Eelco J. Rohling & Heiko Pälike

Southampton Oceanography Centre, Southampton SO14 3ZH, UK

The extent of climate variability during the current interglacial period, the Holocene, is still debated. Temperature records derived from central Greenland ice cores show one significant temperature anomaly between 8,200 and 8,100 years ago, which is often attributed to a meltwater outflow into the North Atlantic Ocean and a slowdown of North Atlantic Deep Water formation—this anomaly provides an opportunity to study such processes with relevance to present-day freshening of the North Atlantic. Anomalies in climate proxy records from locations around the globe are often correlated with this sharp event in Greenland. But the anomalies in many of these records span 400 to 600 years, start from about 8,600 years ago and form part of a repeating pattern within the Holocene. More sudden climate changes around 8,200 years ago appear superimposed on this longer-term cooling. The compounded nature of the signals implies that far-field climate anomalies around 8,200 years ago cannot be used in a straightforward manner to assess the impact of a slowdown of North Atlantic Deep Water formation, and the geographical extent of the rapid cooling event 8,200 years ago remains to be determined.

One notable cooling event between 8.25 and 8.15 thousand years before present (kyr BP) disrupts the remarkable stability of the Holocene epoch in temperature proxy data (ice $\delta^{18}\text{O}$) from the GISP2 Greenland ice core. The age models of the GRIP and North-GRIP Greenland ice cores place it at 8.15–8.05 kyr BP (Fig. 1, and Supplementary Information). The generally accepted explanation for this so-called ‘8.2-kyr-BP event’ envisages curtailment of North Atlantic Deep Water (NADW) formation and its associated northward heat transport, due to a catastrophic meltwater release into the North Atlantic^{1,2}. This notion was supported by dating of a final outburst drainage from glacial lakes Agassiz and Ojibway during the terminal demise of the Laurentide ice sheet, around 8.47 kyr BP (ref. 3). Because the 8.2-kyr-BP event occurred within the Holocene, it might hold clues about the potential response of NADW formation to the current freshening of the North Atlantic⁴. It is also unclear whether the 8.2-kyr-BP event was indeed associated with a significant NADW slowdown^{5,6}, which is thought to be a key mechanism for widespread climate impacts^{7,8}. Clearly, a detailed assessment of the events around 8.2 kyr BP is essential.

Anomalies have been observed around 8.2 kyr BP in palaeoclimate archives on a near-global scale, except for the high southern latitudes^{2,9–11}. Climate models forced with a strong freshwater pulse into the North Atlantic (leading to NADW slowdown) do suggest widespread consequences^{7,8}. On that basis, it seems reasonable to view almost all anomalies observed in climate proxy records at roughly 8 kyr BP as evidence of this scenario for global climate change, ascribing any dating and duration differences to poor age control and resolution in other archives relative to Greenland ice cores². Here we evaluate a selection of well-dated, highly resolved climate proxy records that have been developed since the initial reports. This evaluation bears directly upon: (1) the inferred cause for the 8.2-kyr-BP event; (2) the broader climatic context of the event; and (3) the global distribution of the event.

Anomalies in climate proxy records from Greenland

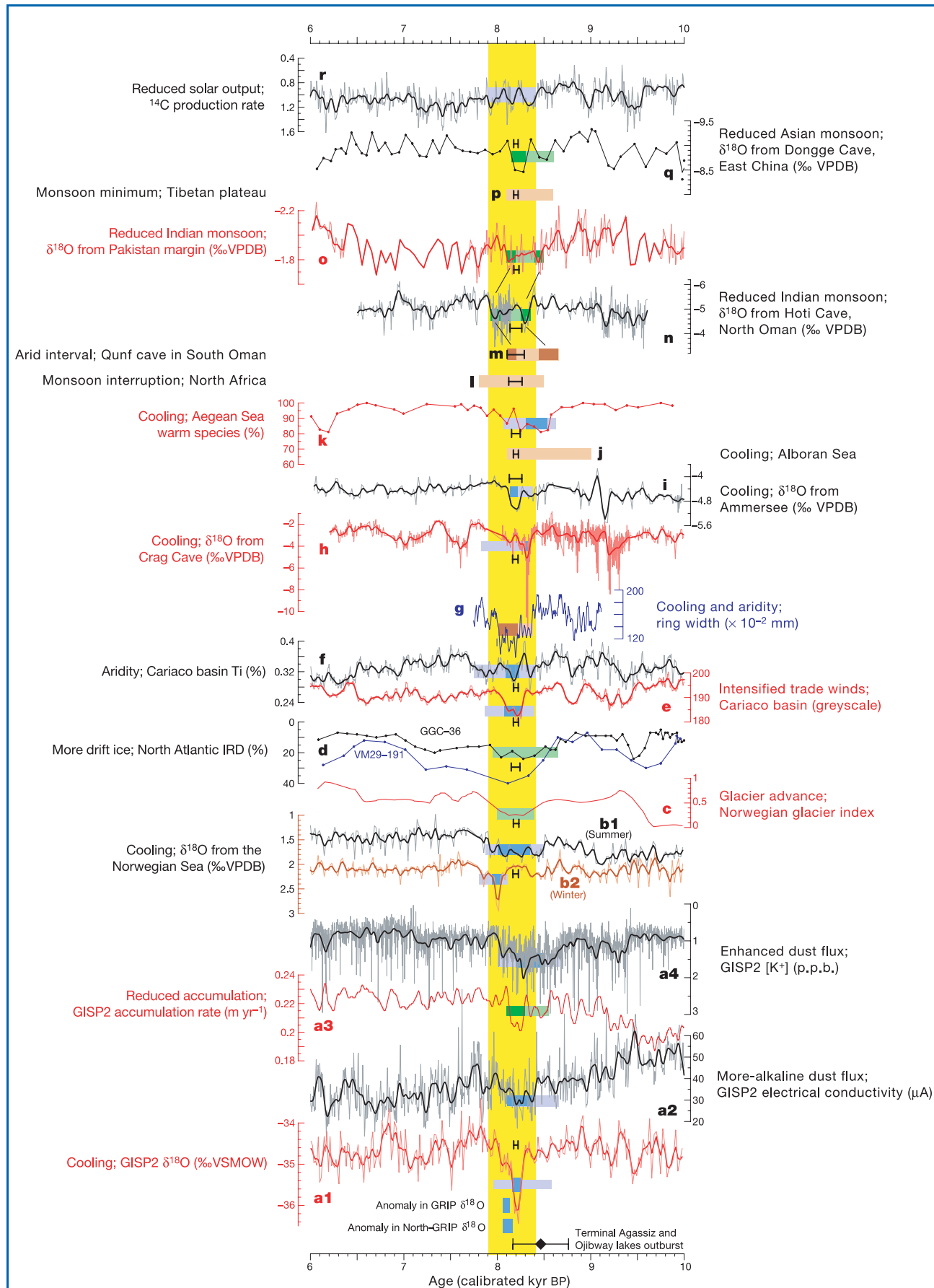
Sharp anomalies centred on about 8.2 kyr BP were found not only in ice $\delta^{18}\text{O}$ (Fig. 1a1), but also in other proxies (for example, chloride) from the Greenland ice cores². However, other Greenland ice-core proxies show anomalies over a much longer time interval. This is best illustrated by the GISP2 potassium series, a proxy for dust supply to Greenland, which contains a broad maximum between

8.65 and 8.00 kyr BP (refs 9, 12; Fig. 1a4). All time intervals cited here are detailed in Supplementary Tables S1 and S2. GISP2 potassium maxima have been interpreted in terms of expansion and intensification of the atmospheric polar vortex^{9,12}. Another proxy for alkaline dust supply to Greenland, electrical conductivity (ECM), also shows a compound maximum between 8.65 and 8.10 kyr BP, with a peak at 8.4–8.1 kyr BP (refs 13, 14) (Fig. 1a2). The GISP2 accumulation rate visually suggests a broad reduction between 8.55 and 8.1 kyr BP with a minimum at 8.3–8.1 kyr BP (Fig. 1a3), although this is not statistically significant (see Supplementary Information). These different proxies, co-registered in a single archive, clearly suggest that the sharp 8.2-kyr-BP event punctuated a longer-term background anomaly. The sharp peak is prominent in the $\delta^{18}\text{O}$ proxy for temperature at the Greenland summit, and the broader anomaly is apparent in proxies for the structure/intensity of atmospheric circulation.

Anomalies in climate proxy records around the North Atlantic

Throughout this region, the sharp 8.2-kyr-BP event, where present, again occurs within a broader anomaly (Fig. 1b–j). A particularly revealing case is portrayed by the anomalies in records b1 and b2 from a Norwegian Sea sediment core¹⁵ (Fig. 1b1,b2). The typical sharp signature of the 8.2-kyr-BP event is found at 8.0–7.9 kyr BP in the $\delta^{18}\text{O}$ record of the left-coiling (sinistral) planktonic foraminifer *Neogloboquadrina pachyderma*, whereas a broad maximum between 8.5 and 7.9 kyr BP is revealed in the $\delta^{18}\text{O}$ of right-coiling (dextral) *N. pachyderma*¹⁵. These records are based on the same sample series, so there can be no argument about the different phasing and durations of the anomalies. The combination of both sharp and broad anomalies within a single archive offers sound evidence that the sharp 8.2-kyr-BP event occurred within a broad climate anomaly between about 8.6/8.5 and 8.0 kyr BP, similar to the pattern highlighted above for the GISP2 ice core. The Norwegian Sea record furthermore offers vital insight into the season-specific nature of the two anomalies, because the $\delta^{18}\text{O}$ record of sinistral *N. pachyderma* may reflect winter temperature, whereas that of dextral *N. pachyderma* may reflect summer temperature¹⁵.

Stalagmite $\delta^{18}\text{O}$ data from the Irish Crag Cave^{16,17} and carbonate $\delta^{18}\text{O}$ data from the German lake Ammersee¹⁸ show a dominant sharp peak anomaly, at 8.35–8.3 and 8.2–8.15 kyr BP, respectively (Fig. 1h–i). These anomalies are (partly) explained in terms of enhanced isotopic fractionation in precipitation during the cold



8.2-kyr-BP event, in agreement with the strong and abrupt cooling seen over the North Atlantic region in climate models that simulate the impact of a meltwater flood^{7,8}. Given that the Norwegian Sea records (Fig. 1b) identify the sharp cooling event as a winter phenomenon, it follows that the Crag Cave and Ammersee records are dominated by isotopic changes in winter precipitation.

The German tree ring width record shows a distinct low between 8.4 and 8.0 kyr BP, culminating in three sharp minima between 8.2 and 8.0 kyr BP (Fig. 1g), caused by colder and more arid climate conditions over central Europe¹⁹. This would agree with the broad period of reduced summer temperatures and superimposed sharp minimum winter temperatures in the Norwegian Sea. Annual temperatures at the Estonian lake Rõuge were reduced between about 8.5 and 8.1 kyr BP (culminating at 8.25–8.15 kyr BP), and erosion-related inorganic sedimentation was increased between 8.5 and 8.1 kyr BP with two sharp maxima at 8.3 and 8.2 kyr BP (ref. 20). These patterns mimic those in the tree-ring record (Fig. 1g), and agree with the substantial 'Finse' advance of Norwegian glaciers at 8.4–8.0 kyr BP (refs 21, 22; Fig. 1c). The Estonian cooling is marked by a decline in *Alnus* pollen, an indication of frost damage in early spring²⁰, which supports the inferred reduction of summer temperature after about 8.5 kyr BP in the Norwegian Sea. Reduced or seasonally delayed summer warming may also (partly) explain the glacier expansion, and would have facilitated the survival and dispersal of sea ice, which governed the Holocene ice-rafted debris (IRD) deposition. Although the initial IRD records showed only a broad maximum^{23,24} (Fig. 1d), records of improved resolution have since revealed the culmination in a sharp IRD spike at the 8.2-kyr-BP event proper²⁵.

If there was an underlying climatic deterioration between about 8.5 and 8.0 kyr BP that was punctuated by the sharp 8.2-kyr-BP event, then we need to be very careful about attributing any anomaly around 8 kyr BP to the 8.2-kyr-BP event itself (as in refs 2, 7 and 17). It would instead be possible that the far-field correlations reflect the underlying climate anomaly, whereas the sharp signals triggered by the meltwater outburst were restricted to a much smaller (for

example, North Atlantic) realm. Archives from remote locations would then be likely to show well-dated anomalies that span four to six centuries, rather than a sharp event over about one century. Because monsoons are the most efficient means for interhemispheric vapour/latent heat transport, we pay specific attention to low- to mid-latitude records for evidence of monsoon variability. Given that the sharp 8.2-kyr-BP event may be most evident in winter-dominated proxies, and the broad underlying anomaly more characteristic of summer conditions, the seasonal significance of proxy records must be taken into account.

Anomalies in climate proxy records from monsoon regions

In the Cariaco basin, a broad multi-peak low in Ti percentages between 8.4 and 7.75 kyr BP reflects aridity due to reduced northward migration of the Intertropical Convergence Zone (ITCZ) in summer²⁶ (Fig. 1f). In contrast, a single sharp anomaly at 8.25–8.1 kyr BP in the greyscale record reflects enhanced intensities of the trade winds, which affect this region in winter²⁷ (Fig. 1e). This combination again portrays the different seasonal responses, with a broad anomaly in a summer-dominated proxy and a sharp event in a winter-dominated proxy.

A well-dated census of African summer-monsoon-fed lakes²⁸ illustrates a well-defined reduction in monsoon intensity or penetration between about 8.5 and 7.8 kyr BP (Fig. 1l) (a similar monsoon reduction exists within the previous (Eemian) interglaciation^{29,30}). A broad Holocene monsoon minimum is also found in the Indian and Asian domains (Fig. 1m–q). Although strong high-frequency variability impedes statistical distinction of the centennial-scale signals in each record individually, records of Indian summer-monsoon intensity from around the Arabian Sea (Fig. 1m–o) show considerable visual similarity to one another^{31–33}, as well as with a marine $\delta^{18}\text{O}$ record from offshore Socotra Island³⁴. Together, these records portray a broad reduction of Indian summer-monsoon activity between roughly 8.5 and 8.0 kyr BP, which both started and ended with a century-scale peak minimum (correlation lines in Fig. 1). There is a great lack of highly resolved

Figure 1 Collection of well-dated climate proxy records used to identify anomalies around 8.2 kyr BP. All records, including those from the ice cores, are presented versus 'calibrated age' in kyr BP, where 'before present' refers to before 1950. The records originate from ice cores, from laminated (usually anoxic) marine sediments, from marine sediments from other anoxic and/or high-accumulation settings that lack significant bioturbation, from cave stalactites or stalagmites and palaeo-lake deposits, and from a peat bog. All are all well dated using a variety of techniques from lamina/layer/varve counting to accelerator mass spectrometric (AMS) ^{14}C dating (including wiggle-matching), and U/Th dating. Error bars with the records portray the size of the one-sided 1σ uncertainty reported for the datings (see Supplementary Information Table S2). All smoothed records shown are based on a 50-yr moving gaussian filter. Coloured blocks identify the anomalies recognized in the various records: blue is statistically significant in digital records (see Supplementary Information); green is visually identified in digital records (not statistically significant); brown is estimated from published graphs. Heavy coloured blocks indicate peak events; lighter blocks highlight broader anomalies. The yellow band marks an interval spanning three conspicuous maxima in ^{14}C production. The black diamond with error bar alongside the age axis indicates the best age estimate for the meltwater flood from lakes Agassiz and Ojibway with 1σ bounds³. Axis labels identify the nature of the anomalies around 8 kyr BP.

Records are grouped by archive; each letter is a separate archive, and subnumbering refers to different proxies co-registered within that archive. Records **a1** to **a4** are all measured from the Greenland Ice Sheet Project II (GISP2) ice core: **a1** is the ice- $\delta^{18}\text{O}$ series⁴⁵, **a2** is the electrical conductivity series^{13,14}, **a3** is a 50-yr smoothing of the ice-accumulation data set⁴⁶ (actual data not shown to avoid clutter), and **a4** is the GISP2 potassium ion series^{9,12} (spikes in actual data clipped at axis maximum value to avoid clutter). Also indicated are the timings of the sharp

$\delta^{18}\text{O}$ anomalies in the Greenland NorthGRIP and GRIP ice cores, which deviate slightly from that in GISP2 (**a1**) owing to small differences between the age models of the ice cores^{47,48}. Records **b1** and **b2** are co-registered in a sediment core from the Norwegian Sea¹⁵: **b1** is $\delta^{18}\text{O}$ of right-coiling *Neogloboquadrina pachyderma*, **b2** is $\delta^{18}\text{O}$ of left-coiling *N. pachyderma*. Record **c** is a Norwegian glacier expansion index based on glacial lake cores investigated at a resolution better than 50 yr (Jostedalssbreen area) (ref. 21, and see also ref. 11). The two records in **d** show IRD percentages in North Atlantic sediment cores VM29-191 (54° N, 15° E) and GGC-36 (45° N, 45° E) (ref. 24). Records **e** and **f** are both from laminated sediment cores from the Cariaco basin: **e** shows greyscale measurements²⁷, and **f** shows titanium concentrations measured by core-logging X-ray fluorescence²⁶. Record **g** (digitized) presents the German tree-ring-width series¹⁹. Record **h** is the Crag Cave $\delta^{18}\text{O}$ series^{16,17}, and **i** is the ostracod-based $\delta^{18}\text{O}$ series for Lake Ammersee¹⁸. Block **j** marks an interval of low sea surface temperatures in the Alboran Sea, westernmost Mediterranean⁴⁹, and record **k** does so for the Aegean Sea, northeastern Mediterranean³⁷. Although of relatively low resolution, the Aegean record is kept within this evaluation because it is well dated and fully supported by a highly resolved (but less well-dated) record for the central Mediterranean⁵⁰. Block **l** identifies the well-established, widespread interruption of the North African monsoon maximum²⁸, and **m** identifies an arid interlude with two distinct peaks reflected in stalagmite $\delta^{18}\text{O}$ for Qunf Cave in South Oman³¹. Record **n** is similar to **m**, but for Hoti Cave in North Oman³². Record **o** shows $\delta^{18}\text{O}$ of the surface-dwelling planktonic foraminifer *Globigerinoides ruber* in a laminated sediment core from the Pakistan margin³³. Block **p** identifies a major interval of reduced monsoon activity identified from cellulose³³ in a peat-bog section in East Tibet³⁵, and record **q** is stalagmite $\delta^{18}\text{O}$ for Dongge Cave, East China³⁶. Record **r** shows ^{14}C production rates from a recent deconvolution of the ^{14}C residuals series^{41,42}.

records from the Asian monsoon sector, but available speleothem and peat-bog records^{35,36} suggest that a monsoon reduction also occurred there between roughly 8.6 and 8.1 kyr BP (Fig. 1p, q).

It cannot be excluded that part of the anomalies in the various low- to mid-latitude records might represent atmospheric or oceanic downstream effects of a meltwater pulse into the North Atlantic and its associated sharp cooling event. However, such a mechanism cannot explain the longer durations of the overall anomaly, or the fact that many records show a well-dated onset of climatic deterioration well before the sharp 8.2-kyr-BP event (Fig. 1, Supplementary Table S2). Note that there is a general lack of low- to mid-latitude records that reflect winter conditions, which might be more likely to show the sharp event than the available summer-dominated records, as illustrated by the data from the Cariaco basin (Fig. 1e, f).

Broad underlying climate anomaly

The early Holocene meltwater flood has been dated³ at 8.47 ± 0.3 kyr BP (1σ). The oldest age for the onset of the sharp cooling in the Greenland ice cores is near 8.3 kyr BP. If it is related to an NADW slowdown due to meltwater addition, this constrains the age of the flood to within a couple of decades before 8.3 kyr BP (refs 7, 8), well within the 1σ bounds of the datings³ and in close agreement with the age of 8.33 ± 0.08 kyr BP for the onset of the sharp climate impact in Ireland^{16,17}.

Climate models with a flood-related cold event of 200 yr suggest associated anomalies in other climate parameters that develop more gradually and may reach twice the duration, but none would start significantly before the flood or cold snap⁸. Evaluation of anomalies that lasted 400 yr or more in our compilation of proxy records (Fig. 1, Supplementary Table S2) shows that 90% (70%) started at or before 8.4 (8.5) kyr BP (Supplementary Table S2). This supports the pivotal observation from the co-registered signals in GISP2 and the Norwegian Sea that a widespread, multi-century climatic deterioration had started well before the flood-related cold event.

The independence of the longer-term anomalies from the meltwater flood is further illustrated by the fact that many are not unique within the Holocene. In GISP2, potassium maxima of comparable magnitude to that of 8.65–8.0 kyr BP are present at 6–5 kyr BP, 3.5–2.5 kyr BP, and AD 1400–1900 (the Little Ice Age), and correlative anomalies have been found outside Greenland^{9,11,12,37}. Spurk *et al.*¹⁹ report a recurrence of tree-ring-width minima similar to that of 8.4–8.0 kyr BP (for example, at 6.1–5.8 kyr BP), which they relate to repeated Holocene IRD episodes^{23,24} and interpret in terms of a recurrent cycle of climatic deteriorations whose impact was amplified around 8.2 kyr BP by the meltwater flood.

Signal comparison between IRD records and the cosmogenic ^{14}C and ^{10}Be records suggests that the repeated Holocene climate deteriorations correspond to intervals of reduced solar output²³, corroborating similar inferences made previously^{9,38,39}. Although high-frequency variability in the Earth's magnetic field might have some influence on cosmogenic isotope production⁴⁰, the solar variability hypothesis is gaining increasing support. ^{14}C production rates^{41,42} reveal a conspicuous 500-yr interval with a succession of three broad ^{14}C production maxima (solar output minima) between 8.4 and 7.9 kyr BP (Fig. 1r). There is more than a passing resemblance with the succession of three ^{14}C production peaks between 0.7 and 0.2 kyr BP that spans the Little Ice Age. The interval of 8.4–7.9 kyr BP agrees well with the timing and duration of the broad climate deterioration highlighted here. The internal three-peak structure may be reflected by internal variability within this broad anomaly, supporting previous reports of remarkable peak-to-peak signal similarities between ^{14}C residuals and proxy records^{31,32,34}.

Compound climate signals around 8.2 kyr BP

The listed evidence for a multi-century climate deterioration, with an onset well before the meltwater flood of about 8.3 kyr ago, indicates that it would be erroneous to attribute all anomalies in climate proxy records around 8 kyr BP to the 8.2-kyr-BP event, in an attempt to map the global impacts of a slowdown in NADW production. Proxies for changes in the meridional extent of major atmospheric circulation features (polar vortex, ITCZ) seem more likely to reflect the underlying deterioration of about 8.5–8.0 kyr ago. In addition, this broad anomaly seems especially evident in summer-biased proxies, and the sharp 8.2-kyr-BP event more evident in winter-biased proxies.

The climate deterioration of about 8.5–8.0 kyr BP is part of a repeating pattern of longer-term anomalies during the Holocene, with its most recent manifestation during the Little Ice Age. It seems related to solar output fluctuations. Climate models for evaluation of the impacts of small reductions in solar output indicate notable impacts on the meridional structure of the atmospheric circulation^{43,44}, in broad agreement with the changes inferred from proxy data, but completely independent from any changes in NADW formation.

Our assessment indicates that the actual geographic extent of the sharp meltwater-induced 8.2-kyr-BP event remains to be established outside the circum-North-Atlantic region, which notably requires the development of new (winter-typical) proxy records in low to middle latitudes. This is especially important if climate proxy signals pertaining to the event are to be used to validate models of NADW slowdown, within the context of discussions on potential changes in deepwater formation in a greenhouse future. □

doi:10.1038/nature03421.

- De Vernal, A., Hillaire-Marcel, C., Von Grafenstein, U. & Barber, D. Researchers look for links among paleoclimate events. *Eos* **78**, 247–249 (1997).
- Alley, R. B. *et al.* Holocene climatic instability: A large event 8000–8400 years ago. *Geology* **25**, 482–486 (1997).
- Barber, D. C. *et al.* Forcing of the cold event of 8,200 years ago by catastrophic drainage of Laurentide lakes. *Nature* **400**, 344–348 (1999).
- Curry, R., Dickson, B. & Yashayaev, I. A change in the freshwater balance of the Atlantic Ocean over the past four decades. *Nature* **426**, 826–829 (2003).
- Weaver, A. J. & Hillaire-Marcel, C. Global warming and the next ice age. *Science* **304**, 400–402 (2004).
- Kuijpers, A. *et al.* Late Quaternary sedimentary processes and ocean circulation changes at the Southeast Greenland margin. *Mar. Geol.* **195**, 109–129 (2003).
- Renssen, H., Goosse, H. & Fichet, T. Modelling the effect of freshwater pulses on the early Holocene climate: the influence of high-frequency climate variability. *Paleoceanography* **17**, doi:10.1029/2001PA000649 (2002).
- Bauer, E., Ganopolski, A. & Montoya, M. Simulation of the cold climate event 8200 years ago by meltwater outburst from Lake Agassiz. *Paleoceanography* **19**, doi:10.1029/2004PA001030 (2004).
- O'Brien, S. R. *et al.* Complexity of Holocene climate as reconstructed from a Greenland ice core. *Science* **270**, 1962–1964 (1995).
- DeMenocal, P., Ortiz, J., Guilderson, T. & Sarnthein, M. Abrupt onset and termination of the African humid period: rapid climate responses to gradual insolation forcing. *Quat. Sci. Rev.* **19**, 347–361 (2000).
- Mayewski, P. A. *et al.* Holocene climate variability. *Quat. Res.* **62**, 243–255 (2004).
- Mayewski, P. A. *et al.* Major features and forcing of high latitude northern hemisphere atmospheric circulation using a 110,000-year-long glaciochemical series. *J. Geophys. Res.* **102**, 26345–26366 (1997).
- Taylor, K. C. *et al.* Ice core dating and chemistry by direct current electrical conductivity. *J. Glaciol.* **38**, 325–332 (1992).
- Taylor, K. C. *et al.* The 'flickering switch' of late Pleistocene climate change. *Nature* **361**, 432–436 (1993).
- Risebrobakken, B., Jansen, E., Andersson, C., Mjelde, E. & Hevrøy, K. A high-resolution study of Holocene paleoclimatic and paleoceanographic changes in the Nordic Seas. *Paleoceanography* **18**, doi:10.1029/2002PA000764 (2003).
- McDermott, F., Matthey, D. O. & Hawkesworth, C. Centennial-scale Holocene climate variability revealed by a high-resolution speleothem $\delta^{18}\text{O}$ record from SW Ireland. *Science* **294**, 1328–1331 (2001).
- Baldini, J. U., McDermott, F. & Fairchild, I. J. Structure of the 8200-year cold event revealed by a speleothem trace element record. *Science* **296**, 2203–2206 (2002).
- Von Grafenstein, U., Erlenkeuser, H., Brauer, A., Jouzel, J. & Johnsen, S. J. A mid-European decadal isotope-climate record from 15,500 to 5000 years B.P. *Science* **284**, 1654–1657 (1999).
- Spurk, M., Leuschner, H. H., Baillie, M. G. L., Briffa, K. R. & Friedrich, M. Depositional frequency of German subfossil oaks: climatically and non-climatically induced fluctuations in the Holocene. *Holocene* **12**, 707–715 (2002).
- Veski, S., Seppä, H. & Ojala, A. E. K. Cold event at 8200 yr B.P. recorded in annually laminated lake sediments in eastern Europe. *Geology* **32**, 681–684 (2004).
- Nesje, A., Matthews, J. A., Dahl, S. O., Berrisford, M. S. & Andersson, C. Holocene glacier fluctuations of Flatbreen and winter-precipitation changes in the Jostedal region, western Norway, based on glaciolacustrine sediment records. *Holocene* **11**, 267–280 (2001).

22. Nesje, A., Dahl, S. O. & Bakke, J. Were abrupt Late glacial and early-Holocene climatic changes in northwest Europe linked to freshwater outbursts to the North Atlantic and Arctic Oceans? *Holocene* **14**, 299–310 (2004).
23. Bond, G. *et al.* Persistent solar influence on north Atlantic climate during the Holocene. *Science* **294**, 2130–2136 (2001).
24. Bond, G. *et al.* A pervasive millennial-scale cycle in North Atlantic Holocene and Glacial climates. *Science* **278**, 1257–1266 (1997).
25. Moros, M. *et al.* Sea surface temperatures and ice rafting in the Holocene North Atlantic: climate influences on northern Europe and Greenland. *Quat. Sci. Rev.* **23**, 2113–2126 (2004).
26. Haug, G. H., Hughen, K. A., Peterson, L. C., Sigman, D. M. & Röhl, U. Southward migration of the Intertropical Convergence Zone through the Holocene. *Science* **293**, 1304–1308 (2001).
27. Hughen, K. A., Overpeck, J. T., Peterson, L. C. & Trumbore, S. Rapid climate changes in the tropical Atlantic region during the last deglaciation. *Nature* **380**, 51–54 (1996).
28. Gasse, F. Hydrological changes in the African tropics since the last glacial maximum. *Quat. Sci. Rev.* **19**, 189–211 (2000).
29. Rohling, E. J. *et al.* African monsoon variability during the previous interglacial maximum. *Earth Planet. Sci. Lett.* **202**, 61–75 (2002).
30. Rohling, E. J. *et al.* Reconstructing past planktic foraminiferal habitats using stable isotope data: a case history for Mediterranean sapropel S5. *Mar. Micropaleontol.* **50**, 89–123 (2004).
31. Fleitmann, D. *et al.* Holocene forcing of the Indian monsoon recorded in a stalagmite from Southern Oman. *Science* **300**, 1737–1739 (2003).
32. Neff, U. *et al.* Strong coherence between solar variability and the monsoon in Oman between 9 and 6 kyr ago. *Nature* **411**, 290–293 (2001).
33. Staubwasser, M., Sirocko, F., Grootes, P. M. & Erlenkeuser, H. South Asian monsoon climate change and radiocarbon in the Arabian Sea during early and middle Holocene. *Paleoceanography* **17**, doi:10.1029/2000PA000608 (2002).
34. Jung, S. J. A., Davies, G. R., Ganssen, G. & Kroon, D. Decadal-centennial scale monsoon variations in the Arabian Sea during the early Holocene. *Geochim. Geophys. Geosyst.* **3**, doi:10.1029/2002GC000348 (2002).
35. Hong, Y. T. *et al.* Correlation between Indian Ocean summer monsoon and North Atlantic climate during the Holocene. *Earth Planet. Sci. Lett.* **211**, 371–380 (2003).
36. Yuan, D. *et al.* Timing, duration, and transitions of the Last Interglacial Asian monsoon. *Science* **304**, 575–578 (2004).
37. Rohling, E. J., Mayewski, P. A., Hayes, A., Abu-Zied, R. H. & Casford, J. S. L. Holocene atmosphere-ocean interactions: records from Greenland and the Aegean Sea. *Clim. Dyn.* **18**, 587–593 (2002).
38. Denton, G. H. & Karlen, W. Holocene climatic variations—Their pattern and possible cause. *Quat. Res.* **3**, 155–205 (1973).
39. Van Geel, B. *et al.* The role of solar forcing upon climate change. *Quat. Sci. Rev.* **18**, 331–338 (1999).
40. St-Onge, G., Stoner, J. S. & Hillaire-Marcel, C. Holocene paleomagnetic records from the St. Lawrence estuary, eastern Canada: centennial- to millennial-scale geomagnetic modulation of cosmogenic isotopes. *Earth Planet. Sci. Lett.* **209**, 113–130 (2003).
41. Muscheler, R., Beer, J. & Vonmoos, M. Causes and timing of the 8200 yr BP event inferred from the comparison of the GRIP ^{10}Be and the tree ring $\Delta^{14}\text{C}$ record. *Quat. Sci. Rev.* **23**, 2101–2111 (2004).
42. Muscheler, R., Beer, J., Kubik, P. W. & Synal, H. A. Geomagnetic field intensity during the last 60,000 years based on ^{10}Be and ^{36}Cl from the Summit ice cores and ^{14}C . *Quat. Sci. Rev.* (in the press).
43. Haigh, J. D. The impact of solar variability on climate. *Science* **272**, 981–984 (1996).
44. Shindell, D., Rind, D., Balachandran, N., Lean, J. & Loneragan, P. Solar cycle variability, ozone, and climate. *Science* **284**, 305–308 (1999).
45. Grootes, P. M., Stuiver, M., White, J. W. C., Johnsen, S. J. & Jouzel, J. Comparison of oxygen isotope records from the GISP2 and GRIP Greenland ice cores. *Nature* **366**, 552–554 (1993).
46. Cuffey, K. M. & Clow, G. D. Temperature, accumulation, and ice sheet elevation in central Greenland through the last deglacial transition. *J. Geophys. Res.* **102**, 26383–26396 (1997).
47. NGRIP members. High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature* **431**, 147–151 (2004).
48. Johnsen, S. *et al.* Irregular glacial interstadials recorded in a new Greenland ice core. *Nature* **359**, 311–313 (1992).
49. Cacho, I., Grimalt, J. O. & Canals, M. Response of the Western Mediterranean Sea to rapid climatic variability during the last 50,000 years: a molecular biomarker approach. *J. Mar. Syst.* **33–34**, 253–272 (2002).
50. Sprovieri, R., Di Stefano, E., Incarbona, A. & Gargano, M. E. A high-resolution record of the last deglaciation in the Sicily Channel based on foraminifera and calcareous nannofossil quantitative distribution. *Paleoceanogr. Palaeoclimatol. Palaeoecol.* **202**, 119–142 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank all colleagues who generously made their data available, either directly or through the NOAA National Geophysical Data Centre, and J. Thomson for discussions and suggestions.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to E.J.R. (E.Rohling@soc.soton.ac.uk).

The genome sequence of the rice blast fungus *Magnaporthe grisea*

Ralph A. Dean¹, Nicholas J. Talbot², Daniel J. Ebbole³, Mark L. Farman⁴, Thomas K. Mitchell¹, Marc J. Orbach⁵, Michael Thon³, Resham Kulkarni^{1*}, Jin-Rong Xu⁶, Huaqin Pan¹, Nick D. Read⁷, Yong-Hwan Lee⁸, Ignazio Carbone¹, Doug Brown¹, Yeon Yee Oh¹, Nicole Donofrio¹, Jun Seop Jeong¹, Darren M. Soanes², Slavica Djonovic³, Elena Kolomiets³, Cathryn Rehmeyer⁴, Weixi Li⁴, Michael Harding⁵, Soonok Kim⁸, Marc-Henri Lebrun⁹, Heidi Bohnert⁹, Sean Coughlan¹⁰, Jonathan Butler¹¹, Sarah Calvo¹¹, Li-Jun Ma¹¹, Robert Nicol¹¹, Seth Purcell¹¹, Chad Nusbaum¹¹, James E. Galagan¹¹ & Bruce W. Birren¹¹

¹Center for Integrated Fungal Research, North Carolina State University, Raleigh, North Carolina 27695, USA

²School of Biological and Chemical Sciences, University of Exeter, Washington Singer Laboratories, Exeter EX4 4QG, UK

³Department of Plant Pathology and Microbiology, Texas A&M University, College Station, Texas 77843, USA

⁴Department of Plant Pathology, University of Kentucky, Lexington, Kentucky 40546, USA

⁵Department of Plant Pathology, University of Arizona, Tucson, Arizona 85721, USA

⁶Department of Botany and Plant Pathology, Purdue University, West Lafayette, Indiana 47907, USA

⁷Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh EH9 3JH, UK

⁸School of Agricultural Biotechnology, Seoul National University, Seoul 151-742, Korea

⁹FRE2579 CNRS-Bayer, Bayer CropScience, 69263 Lyon Cedex 09, France

¹⁰Agilent Technologies, Wilmington, Delaware 19808, USA

¹¹Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02141, USA

* Present address: RTI International, Research Triangle Park, North Carolina 27709, USA

***Magnaporthe grisea* is the most destructive pathogen of rice worldwide and the principal model organism for elucidating the molecular basis of fungal disease of plants. Here, we report the draft sequence of the *M. grisea* genome. Analysis of the gene set provides an insight into the adaptations required by a fungus to cause disease. The genome encodes a large and diverse set of secreted proteins, including those defined by unusual carbohydrate-binding domains. This fungus also possesses an expanded family of G-protein-coupled receptors, several new virulence-associated genes and large suites of enzymes involved in secondary metabolism. Consistent with a role in fungal pathogenesis, the expression of several of these genes is upregulated during the early stages of infection-related development. The *M. grisea* genome has been subject to invasion and proliferation of active transposable elements, reflecting the clonal nature of this fungus imposed by widespread rice cultivation.**

Outbreaks of rice blast disease are a serious and recurrent problem in all rice-growing regions of the world, and the disease is extremely difficult to control^{1,2}. Rice blast, caused by the fungus *Magnaporthe grisea*, is therefore a significant economic and humanitarian problem. It is estimated that each year enough rice is destroyed by rice blast disease to feed 60 million people³. The life cycle of the rice blast fungus is shown in Fig. 1. Infections occur when fungal spores land and attach themselves to leaves using a special adhesive released from the tip of each spore⁴. The germinating spore develops an appressorium—a specialized infection cell—which generates enormous turgor pressure (up to 8 MPa) that ruptures the leaf cuticle, allowing invasion of the underlying leaf tissue^{5,6}. Subsequent colonization of the leaf produces disease lesions from which the fungus sporulates and spreads to new plants. When rice blast infects young rice seedlings, whole plants often die, whereas spread of the disease to the stems, nodes or panicle of older plants results in nearly total loss of the rice grain². Different host-limited forms of *M. grisea* also infect a broad range of grass species including wheat, barley and millet. Recent reports have shown that the fungus has the capacity to infect plant roots⁷.

Here we present our preliminary analysis of the draft genome sequence of *M. grisea*, which has emerged as a model system for understanding plant–microbe interactions because of both its economic significance and genetic tractability^{1,2}.

Acquisition of the *M. grisea* genome sequence

The genome of a rice pathogenic strain of *M. grisea*, 70-15, was sequenced through a whole-genome shotgun approach. In all, greater than sevenfold sequence coverage was produced, and a

summary of the principal genome sequence data is provided in Table 1 and Supplementary Table S1. The draft genome sequence consists of 2,273 sequence contigs longer than 2 kilobases (kb), ordered and orientated within 159 scaffolds. The total length of all sequence contigs is 38.8 megabases (Mb), and the total length of the scaffolds, including estimated sizes for the gaps, is 40.3 Mb. The genome assembly has high sequence accuracy—96% of the bases have quality scores of greater than 40—and long-range continuity, with 50% of all bases residing in scaffolds longer than 1.6 Mb.

Reconstruction of the *M. grisea* genome was aided by the availability of genome maps^{8,9} (Supplementary Methods S1). Thirty-three scaffolds, representing 32.8 Mb or 85% of the draft assembly, were ordered on the genetic map and assigned to each of the seven chromosomes by virtue of containing an anchored genetic marker. In addition, 19 scaffolds (65% of genome assembly) contained more than one marker and could thus be oriented on the map. The ends of chromosomes were identified by the telomere repeat motif (TTAGGG)_n. Thirteen telomeric sequences were placed at the ends of scaffolds, of which six could be placed at the ends of chromosomes, whereas the remainder were associated with unanchored scaffolds (Supplementary Table S2). Genome coverage was estimated by aligning 28,682 *M. grisea* expressed sequence tags (ESTs), representing genes expressed during a range of developmental stages and environmental conditions^{10,11}. Approximately 94% of the ESTs were aligned to the genome assembly, despite many of these ESTs being from different strains.

The gene content of a plant pathogenic fungus

Within the *M. grisea* genome, 11,109 genes were predicted with

protein products of longer than 100 amino acids (Supplementary Methods S2). These predicted genes comprise 48% of the assembly (1 gene per 3.5 kb). For comparison, 10,082 genes were predicted in *Neurospora crassa*, a related pyrenomycete, and 9,457 were predicted in the more distantly related plectomycete *Aspergillus nidulans* (also known as *Emericella nidulans*; <http://www.broad.mit.edu/annotation/fungi/aspergillus/>). Neither of these species causes plant disease. At the amino acid level, *M. grisea* orthologues in *N. crassa* and *A. nidulans* show an average identity of 47% and 46%, respectively.

Given that the *M. grisea* genome possesses more genes than the non-plant pathogens *N. crassa* and *A. nidulans*, and that these fungi are well-studied model organisms, we compared genes between these species to identify potential cases of gene family expansion that might be associated with the evolution of *M. grisea* as a plant pathogen.

Single linkage clustering of protein sequences resulted in 348 families containing five or more genes. Overall, more proteins from *M. grisea* (1,266, $P < 0.001$) and *A. nidulans* (1,424, $P < 0.001$) were classified into families as compared with those from *N. crassa* (950). An important factor in this comparison is that *N. crassa* has a

process called repeat-induced point mutation (RIP), a mechanism that eliminates duplicated genes during meiosis and therefore limits the ability of *N. crassa* to undergo paralogous gene duplication and consequent gene family expansion¹². Twenty-eight families were found with significant differences ($P < 0.05$) in gene content between the three species, nine of which were larger in *M. grisea* than they were in the genome of either *N. crassa* or *A. nidulans* (Table 2; see also Supplementary Table S3). Several gene families expanded in *M. grisea* exhibited sequence similarity to proteins that are involved in fungal pathogenicity¹³.

Phylogenetic trees (Supplementary Fig. S1) of the expanded gene families in *M. grisea* indicated that none of the families showed evidence for recent lineage-specific expansion. These gene families in *M. grisea* may therefore be the result of ancient gene duplication events followed by loss of gene family members in the *N. crassa* and *A. nidulans* lineages.

Genome architecture and co-linearity

Co-linearity of chromosome segments (synteny) is widely reported for plant and animal genomes¹⁴, but little is known about this for filamentous ascomycete fungi. Analysis of orthologous pairs of genes in *M. grisea* and *N. crassa* revealed no evidence for extensive regions of conserved synteny, although linkage group assignments were often conserved between the two species, notably chromosome 7 in *M. grisea* with linkage group I of *N. crassa* (Supplementary Table S4). Only 113 regions containing four or more genes were found to be co-linear between *M. grisea* and *N. crassa*. One example, also conserved in several other filamentous fungi, is the quinate/shikimate (Qa) metabolic pathway gene cluster (Supplementary Fig. S2). This seven-gene cluster, spanning ~20 kb, which participates in quinate use and aromatic amino acid catabolism, is found on chromosome 3 in *M. grisea*, and is syntenic in *N. crassa*, *A. nidulans* and other filamentous ascomycetes, but is not present in *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* or other yeasts.

M. grisea has a family of novel G-protein-coupled receptors

To be a successful plant pathogen, a fungus must undergo a series of morphological and physiological programmes⁵. During these developmental transitions, the pathogen also overcomes (or suppresses) the plant's innate immune system and perturbs host metabolism and cell signalling to favour fungal growth. We have attempted to define the most important characteristics of the genome of *M. grisea* that are associated with its pathogenic lifestyle.

G-protein-coupled receptors with seven transmembrane helices (GPCRs) transduce environmental signals, by means of heterotrimeric G proteins, to activate secondary messengers and regulate gene expression^{15,16}. The *M. grisea* genome contains a large repertoire of GPCR-like genes, including 61 not previously described. Twelve of these genes form a subfamily (Supplementary Figs S3 and S4), and contain a conserved fungal-specific extracellular membrane-spanning domain (CFEM) at the amino terminus¹⁷ that resembles the epidermal growth factor (EGF) module present in certain human GPCRs¹⁵. Notably, one of the CFEM-GPCRs, Pth11,

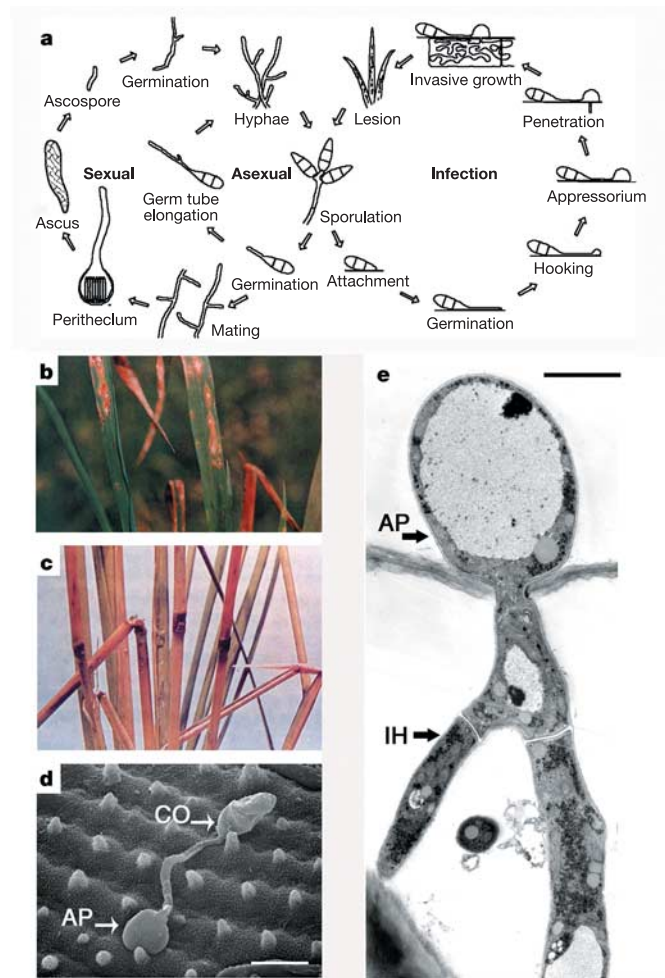


Figure 1 Life history of *Magnaporthe grisea*. **a**, Asexual spores called conidia germinate and develop a specialized infection structure, the appressorium. Invasive growth within and between cells culminates with sporulation and lesion formation. Sexual reproduction occurs when two strains of opposite mating type meet and form a peritheclum in which ascospores develop. Once released, ascospores can develop appressoria and infect host cells. **b**, Rice blast lesions on leaves. **c**, Node blast symptoms on rice, which cause lodging and yield loss. **d**, Scanning electron micrograph of an appressorium (AP) forming after germination of a conidium (CO) on the rice leaf cuticle. The appressorium generates enormous turgor to provide mechanical force to breach the leaf surface. Scale bar, 10 μ m. **e**, Transmission electron micrograph of an appressorium, showing a penetration hypha entering the leaf and developing invasive hyphae (IH). Scale bar, 5 μ m.

Table 1 *Magnaporthe grisea* assembly features

General genome features	Value
Size (bp)	37,878,070
Chromosomes	7
(G+C) percentage	51.6
Protein-coding genes	11,109
tRNA genes	316
Per cent coding	40.5
Average gene size (bp)	1,683
Average intergenic distance (bp)	1,503
Predicted protein-coding sequences	
Conserved hypothetical proteins	8,868 (79%)
Predicted proteins*	2,233 (20%)

*No similarity to known sequences.

is required for appressorium formation and pathogenesis^{2,18}, and *M. grisea* has the largest number of CFEM-GPCR proteins among sequenced fungi. In contrast, only one was detected in *N. crassa* and two in *A. nidulans*, and this type of GPCR is completely absent in the non-filamentous ascomycetes *S. cerevisiae*, *S. pombe*, *Candida albicans* and *Pneumocystis carinii*, or the basidiomycete fungi *Cryptococcus neoformans*, *Ustilago maydis* and *Phanerochaete chrysosporium* (Fig. 2). Moreover, whole-genome microarray analysis confirmed that the identified CFEM-GPCR-like proteins are expressed during infection-related development, and two CFEM-GPCR genes are specifically upregulated when the fungus is undergoing appressorium formation, as shown in Fig. 2. Together, these data suggest that *M. grisea* has greater flexibility to react to extracellular signals compared with saprobic fungi. The pathogenic lifestyle, which involves transitions from plant surface to colonization of distinct plant tissues, may therefore require the ability to respond to a wider range of physical and environmental stimuli.

Virulence-associated signalling pathways in *M. grisea*

The *M. grisea* genome contains three mitogen-activated protein kinase (MAPK) cascades that regulate appressorium development, penetration peg formation and adaptation to hyper-osmotic stress (Fig. 3). Although there are similar numbers of MAPK pathways in other fungi (*N. crassa* has three, *A. nidulans* four and *S. cerevisiae* five), the processes regulated by the pathways in different organisms are distinct. Two of the three MAPK pathways in *M. grisea* control virulence-associated development. The core of the Pmk1 MAPK (pathogenicity MAPK) pathway, which regulates appressorium formation in *M. grisea*, is clearly related to both the pheromone signalling and filamentation pathways from *S. cerevisiae*, and Pmk1 is able to function in place of either the Fus3 or Kss1 MAPK in yeast¹⁹. In *M. grisea*, however, Pmk1 pathway components such as the MAPK are involved in mating and pathogenesis, whereas Mst12, the Ste12-related transcription factor, is dispensable for mating altogether. These observations demonstrate the functional divergence of these signalling pathways. Furthermore, homologues of Pmk1 are required for fungal virulence in all plant pathogens in which they have been investigated²⁰. The Pmk1 signalling pathway revealed from the genome sequence is therefore likely to be of generic importance for fungal pathogenesis.

A principal difference in the operation of MAPK signalling in *M. grisea* compared with *S. cerevisiae* is the absence of a clear Ste5 homologue. A homologue also seems to be absent in other filamentous ascomycetes, including *A. nidulans* and *N. crassa*. In yeast, Ste5 is the scaffold protein that conditions specificity in MAPK signalling²¹. The absence of an identifiable scaffold implies that either a hitherto uncharacterized protein tethers the MAPK signalling module together in *M. grisea* and provides specificity in signal transmission, or that MAPK specificity is governed by a different mechanism in this fungus.

Cyclic AMP signalling is required for the induction of appressorium formation and for the turgor-driven process that leads to plant infection. Until now it has not been completely clear how this cAMP

signal is transmitted. A known cAMP-dependent protein kinase A (PKA) catalytic subunit (CPKA) is required for appressorium maturation, but is dispensable for early appressorium development^{19,22}. The genome has revealed a gene that putatively encodes a second PKA catalytic subunit (MG02832.4). Evaluation of this gene may shed light on how cAMP-mediated signalling operates both at initiation and later stages of appressorium formation. The Pth11 GPCR also operates upstream of the cAMP signalling pathway in *M. grisea*¹⁸, suggesting that the CFEM-GPCR family could provide a number of distinct and unforeseen inputs into this pathway.

Turgor-driven plant infection by *M. grisea*

Appressoria of *M. grisea* generate the enormous turgor pressure needed to breach the plant cuticle through accumulation of up to 3 M concentrations of glycerol. Analysis of the *M. grisea* genome suggests that germinating spores possess considerable versatility in their capacity to synthesize glycerol in the appressorium from storage products. Notably, in contrast to *S. cerevisiae*, where fatty acid β -oxidation occurs solely in peroxisomes, *M. grisea* has several genes that putatively encode acyl-CoA dehydrogenases, but does not appear to possess a gene encoding acyl-CoA oxidase. Thus, β -oxidation in *M. grisea* may occur both in mitochondria and in catalase-free glyoxysome-like bodies, as in *N. crassa*, which would allow use of a wide range of fatty acid substrates, including branched-chain fatty acids²³. *M. grisea* also seems to have the capacity to synthesize glycerol from the glycolytic intermediates dihydroxyacetone phosphate and dihydroxyacetone. Activity of both NADH-dependent glycerol-3-phosphate dehydrogenase and NADPH-dependent glycerol dehydrogenase has been reported in developing appressoria of *M. grisea*²⁴. Taken together, the apparent flexibility in lipid metabolism and ability to divert intermediates from glycolysis may be important for rapid glycerol accumulation during appressorium development.

M. grisea possesses a complex secreted proteome

The secreted proteome is a crucial component of the ability of fungi to perceive and respond to the environment. We identified the presence of 739 proteins that are predicted to be secreted by *M. grisea*, approximately twice the corresponding number for *N. crassa*. Part of this difference reflects an expansion in protein families in *M. grisea*. For example, 163 putatively secreted proteins occur in families containing at least twice as many members as the corresponding family in *N. crassa* (Supplementary Table S5). Several of these expanded families are predicted to encode enzymes for degradation of the plant cell wall and cuticle. For example, eight genes in *M. grisea* putatively encode cutinases—methyl esterases that degrade cutin, the waxy polymer that forms the leaf cuticle. Several of these genes are significantly upregulated during infection-related development (Fig. 2). Previous experiments involving deletion of the cutinase *CUT1* (MG01943.4) led to the conclusion that enzymatic degradation of the cuticle was dispensable for plant infection²⁵. However, *CUT1* is not among those genes differentially regulated during appressorium formation. Any of the remaining seven

Table 2 Gene family content differences between *M. grisea*, *N. crassa* and *A. nidulans*

ID	Magnaporthe	Neurospora	Aspergillus	Annotation
78†	9	1*	3	Feruloyl esterase B precursor
247	6	1	0*	UDP-glucuronosyltransferase precursor, microsomal
134	8	4	0*	350–400 amino acid polypeptide
108†	8	0*	3	Cutinase precursor
4	34	11*	21	Isotrichodermin C-15 hydroxylase (cytochrome P450)
360	6	0*	0*	~180 amino acid cysteine-rich (6–8 cysteine residues) polypeptide
31†	15	1*	8	Cytochrome P450
96†	11	1*	0*	Subtilisin-like serine protease
180	10	0*	0*	~150 amino acid cysteine-rich (6–10 cysteine residues) polypeptide

Gene families larger in *M. grisea* than in *N. crassa* or *A. nidulans* (significant G-statistic; $P < 0.05$) are shown.

*Values are significantly smaller ($P < 0.05$) in pair-wise post-hoc G-tests of *M. grisea* versus *N. crassa* or *M. grisea* versus *A. nidulans*.

†Families implicated in pathogenicity¹³.

cutinases may be capable of complementing the activity of Cut1 in the $\Delta cut1$ mutant. Coupled with the absence of cutinase-encoding genes in *N. crassa*—which colonizes dead plant tissues—these data strongly suggest that cutinases have a significant role in *M. grisea*.

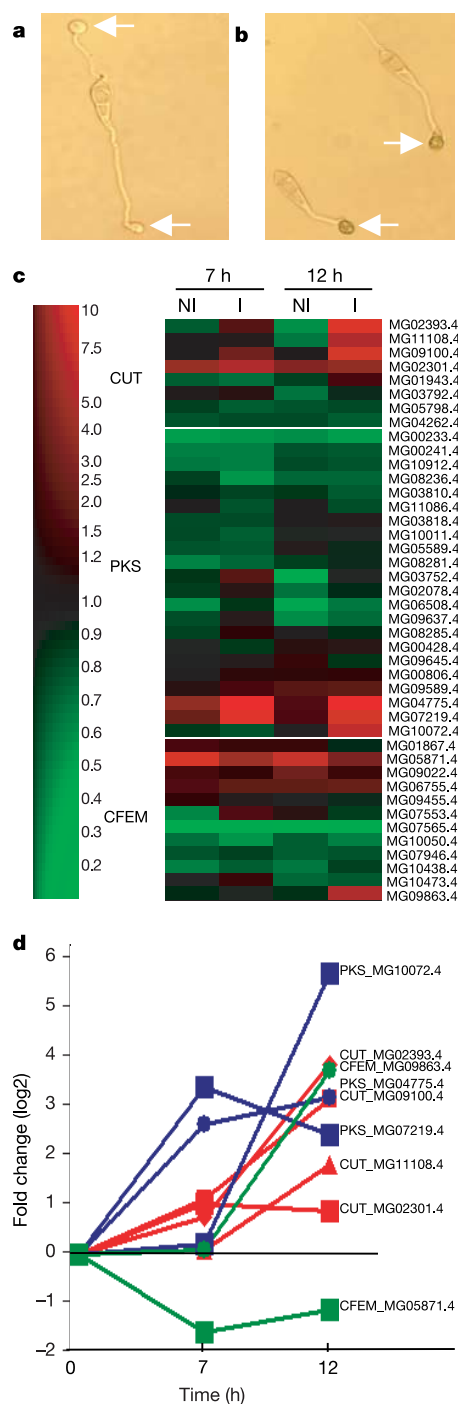


Figure 2 Differential expression of selected *M. grisea* genes during infection-related development. **a, b**, Appressorium formation (arrows) at 7 h (**a**; cessation of polar growth and tip hooking) and 12 h (**b**; tip swelling and melanization) after germination. **c**, Expression profiles of selected genes from conidia germinated on hydrophobic appressorium-inducing (I) and hydrophilic non-appressorium-inducing (NI) surfaces. The colour scale indicates transcript abundance relative to ungerminated conidia: red, increase in transcript abundance; green, decrease in transcript abundance. **d**, Fold change in transcript abundance (cutinases (CUT), red; polyketide synthases (PKS), blue; and Pth11 homologues with CFEM domain, green) at 7 and 12 h after germination on a hydrophobic surface compared with a hydrophilic surface. Only genes exhibiting a twofold or greater (≤ -1 or ≥ 1) change in expression are shown.

Among the secreted proteins predicted for *M. grisea*, many contained consensus carbohydrate substrate-binding domains, consistent with a role in attachment and colonization of plant tissue (Supplementary Table S6). Specifically, a role for chitin-binding proteins in plant–fungus interactions has been proposed from studies of the avirulence protein Avr4 of the tomato pathogen *Cladosporium fulvum*²⁶. Avr4 is a chitin-binding protein that may act to protect the fungal cell wall from chitinases produced by plants' innate immune response. Inspection of proteins containing the cysteine pattern found in Avr4 and other variant motifs revealed a novel pattern that is highly abundant in *M. grisea*. The novel variant cysteine pattern CX₂CCX₂C is present in 36 copies of 21 predicted proteins. In contrast, this pattern occurs just eight times in *A. nidulans*, four times in *N. crassa*, and not at all in *S. cerevisiae* open reading frames (Supplementary Fig. S5). This motif probably represents a variant chitin-binding motif whose abundance may be diagnostic for plant-associated filamentous ascomycetes.

Fungal effectors and PAMPs

Pathogenic microorganisms of plants are known to secrete proteins directly into host plant cells to perturb host cell signalling or suppress the plant innate immune system^{27,28}. The plant adaptive immune system has, in turn, evolved to recognize pathogen effector proteins (often termed pathogen-associated molecular patterns, PAMPs). In the *M. grisea*–rice interaction, this immunity is governed by a gene-for-gene system (one gene in the host conditions resistance to a pathogen effector protein encoded by a single pathogen gene).

Interrogation of the *M. grisea* genome for putative effector proteins revealed three families of putatively secreted, cysteine-rich polypeptides (clusters 180, 360 and 641) and a protein family with similarity to the necrosis-inducing peptide of *Phytophthora infestans* (NPP1, pfam05630)²⁹. As described above, cysteine-rich polypeptides are recognized as PAMPs, as exemplified by the *C. fulvum*–tomato interaction³⁰. In addition, a novel family of proteins with similarity to the N-terminal half (~150 amino acids) of the enterotoxin A chain (pfam01375) was noted. This region of the enterotoxin A chain contains ADP-ribosylation activity³¹, which suggests that the *M. grisea* protein may interact with rice GTP-binding proteins. The genome of the sequenced *M. grisea* strain 70-15 contains four known *M. grisea* avirulence genes: AVR-Pita, ACE1, PWL2 and PWL3. No orthologues were found to *M. grisea* PWL1, PWL4 or AVR1-CO39, or to well-characterized AVR genes from other pathogenic fungi including Avr2, Avr4, Avr9, ECP2, ECP3 and ECP5 from *C. fulvum*³⁰, and NIP1 from *Rhynchosporium secalis*³². This highlights the diversity and lack of sequence similarity or conservation in the fungal avirulence gene products identified so far.

Secondary metabolic pathways of *M. grisea*

Filamentous fungi are well known producers of secondary metabolites, which in nature fulfil a variety of functions probably to allow for niche exploitation. Plant pathogenic fungi produce diverse secondary metabolites that aid in pathogenicity, such as host-selective toxins³³. The *M. grisea* genome displays a considerable capacity for production of secondary metabolites. There are 23 genes predicted to encode polyketide synthases (PKS), compared with seven PKS genes in *Neurospora*, and three of the PKS-encoding genes in particular are upregulated during infection-related development (Fig. 2). Beyond containing the ketosynthase domain, the structure of these proteins is highly divergent (Supplementary Fig. S6). Most of the 23 PKS genes appear to occur in gene clusters with neighbouring genes that are predicted to encode enzymes such as cytochrome P450 and mono-oxygenases, which typically modify or customize the polyketide backbone to form a functional secondary metabolite (Supplementary Table S7). The diversity of PKS genes in filamentous ascomycete fungi and the poor conservation of clearly orthologous PKS genes, even in related species³⁴, indicates

that considerable variability is likely to exist in the polyketide metabolites generated by pathogenic fungi.

Non-ribosomal peptide synthetases (NRPS), which catalyse production of cyclic peptides including numerous toxins, also seem to be well represented in the *M. grisea* genome. Two distinct subclasses exist: those that exist separately and those that are fused to a PKS (PKS–NRPS). Overall, there are six likely NRPS genes and eight PKS–NRPS genes in the *M. grisea* genome: six full-length PKS–NRPS genes and two PKS–NRPS genes with a truncated NRPS domain. This contrasts with two predicted NRPS genes and one NRPS-related gene reported in the *N. crassa* genome sequence. One of the hybrid PKS–NRPS proteins is encoded by the *ACE1* gene, which has recently been shown to act as an avirulence gene³⁵. The large number, and expression profile, of PKS and NRPS genes in *M. grisea* is consistent with the requirements of a fungal pathogen in adapting to diverse environments, perturbing host metabolism and ultimately causing plant cell death.

Repetitive DNA and repeat-induced point mutations

M. grisea exhibits a high degree of genetic variability, and novel pathogenic variants capable of infecting formerly resistant host plants arise with alarming frequency during rice cultivation³⁶. Such

gains of virulence are often associated with transposon-mediated inactivation or deletion of PAMP-encoding genes whose products trigger the plant adaptive immune system^{37,38}. Thus, an understanding of the natural history of repetitive elements in *M. grisea* not only provides an insight into their impact on genome evolution but also sheds light on mechanisms of pathogenic variation. Approximately 9.7% of the *M. grisea* genome assembly comprises repetitive DNA sequences longer than 200 base pairs (bp) and with greater than 65% similarity. Four previously unknown repeats were discovered, as were alternative forms for three previously described transposons. The genome sequence also revealed full-length sequences of two elements for which only incomplete sequences were previously available.

Most repetitive sequences in the assembly are retrotransposons comprising eight major families (Table 3). Five are retroelement families and three are DNA transposons. These repetitive elements are not uniformly distributed in the genome assembly, but form discrete clusters (Supplementary Fig. S7 and Supplementary Note S1)³⁹. Further examination revealed many examples of elements inserted into copies of themselves or other elements. Examination of integration events occurring within other transposons provides evidence for extensive past recombination in the genome

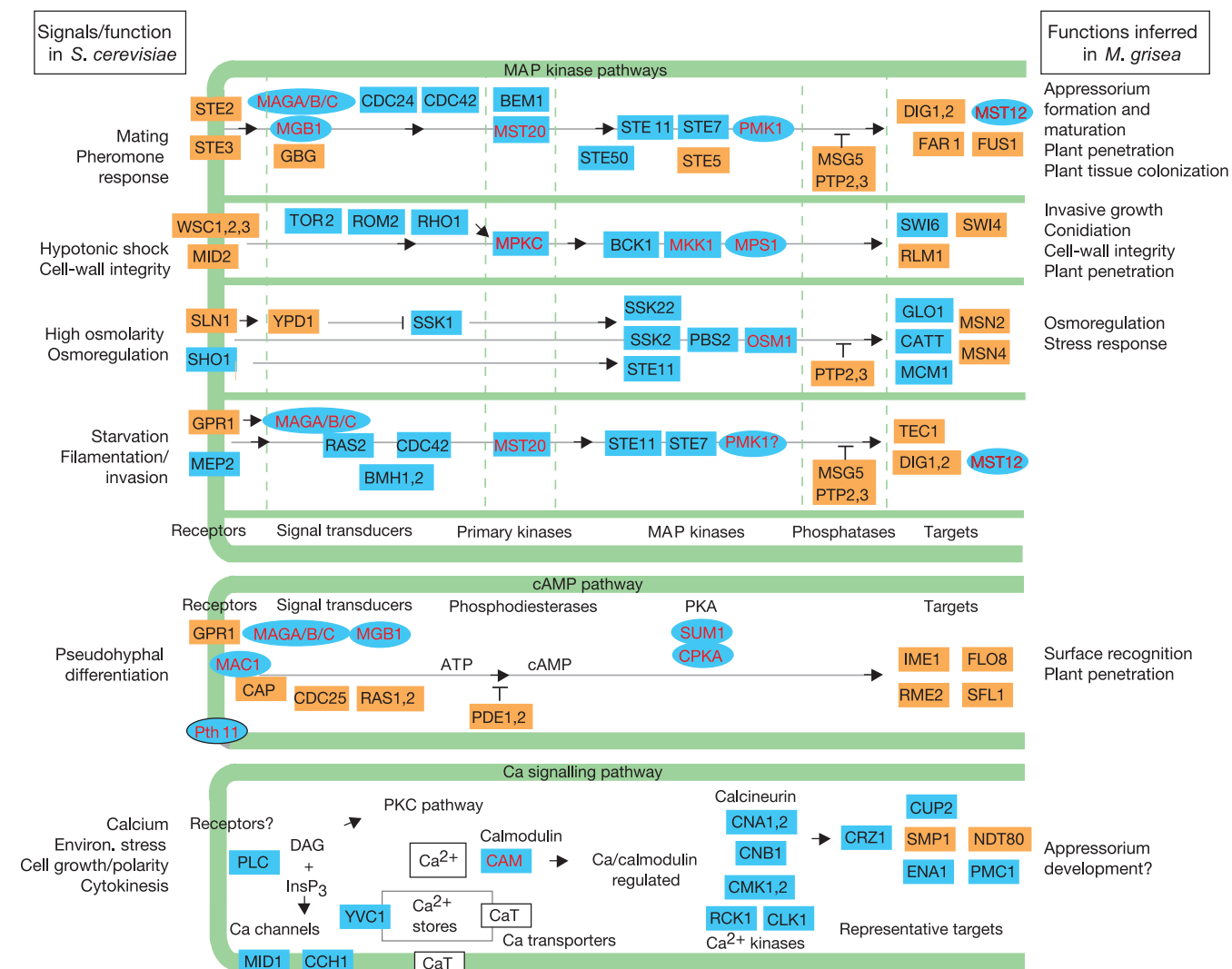


Figure 3 Comparison of major signalling pathways between *S. cerevisiae* and *M. grisea*. The core components of MAP kinase-, cAMP-dependent and calcium-signalling pathways are conserved; however, receptors and downstream targets are less conserved. Protein names in red indicate previously identified *M. grisea* homologues; names in black are *S. cerevisiae* protein names. The colour of the boxes around the protein names

indicates the degree of conservation between *M. grisea* and *S. cerevisiae* proteins based on BLASTP: blue, e -value $< 1 \times 10^{-10}$; orange, e -value $> 1 \times 10^{-10}$. Ovals indicate instances where the *M. grisea* homologue was shown to be required for pathogenicity. Ovals with black borders are *M. grisea* proteins not found in *S. cerevisiae* or not implicated in the *S. cerevisiae* pathway. InsP₃, inositol triphosphate; DAG, diacylglycerol.

(Supplementary Discussion S1). Given the prevalence of repetitive elements and their ability to participate in recombination, it is perhaps surprising that an organism could tolerate such substantial genomic change. However, in nature, rice pathogenic strains of *M. grisea* propagate asexually and, as such, genome organization is rarely, if ever, subject to the potentially catastrophic effects of meiotic recombination involving homologous chromosomes with radically different structures⁴⁰. Thus, rearrangements that would normally have been purged by meiosis appear to have been maintained in the absence of deleterious effects on vegetative fitness. Some rearrangements are expected to have positive fitness benefits, especially those that result in loss of genes whose products would normally trigger defence responses in potential hosts. Many host-specificity genes in *M. grisea* are situated in transposon-rich regions of the genome; this arrangement provides ample opportunity for host-range expansion through gene loss^{37,41}.

The prevalence of intact and essentially identical repeated DNA elements in the genome suggests that *M. grisea* has been unable to stop their proliferation. This is of interest, because RIP has previously been reported to occur in at least one strain (Br48) of *M. grisea*⁴². Inspection of *M. grisea* repeats reveals evidence for RIP (Supplementary Methods S3). However, for Pyret, the repeat family showing the most extensive RIP-like mutations, only three-quarters exhibit signs of RIP and show an average of only 11.4% sequence divergence from the reference sequence. Other elements such as Pot2 are present in over 100 apparently intact copies, displaying an average of 99.4% nucleotide identity. In contrast, all repeat elements in *N. crassa* show heavy RIP mutation, typically to greater than 20% divergence¹². The presence of many intact, highly similar elements despite evidence of RIP can be explained by a number of factors. First, it is possible that RIP was lost in the sequenced strain.

Table 3 Characterization of repeat elements in the *M. grisea* genome

Repeat name	Length (bp)	Per cent of genome*	IR/LTR length (bp)	Evidence of RIP†
Inverted repeat DNA transposons				
Pot2/4		1.3§		—§
Pot2‡	1,857		45	
Pot4	1,858		42	
Pot3	1,860	0.31	42	—
Occan	2,686	0.33	74	++
LTR retrotransposons				
MAGGY	5,638	0.86	253	+
MGLR-3	6,304	0.10	213	++
Pyret‡		1.47§		+++§
Pyret-1¶	7,407		450	
Pyret-2	7,441		474	
Pyret-3	7,407		476	
RETRO5	7,623	0.48	537	—
RETRO6‡		0.37§		+++§
RETRO6-1	5,871		193	
RETRO6-2	5,905		194	
RETRO7‡		0.49§		+++§
RETRO7-1	6,210		216	
RETRO7-2	6,134		208	
Non-LTR retrotransposons (LINEs and SINEs)				
MGL	5,977	1.43	—	++
Mg-SINE-A	472	0.16	—	++
REPBUF	1,111	0.01	—	—
Miscellaneous				
SR2	175	0.01	—	+
SR3	88	0.01	—	—
5S rRNA	70	0.01	—	—

Shown are elements that have a copy number of ten or greater and for which at least one full-length copy exists in the genome. IR, inverted repeat; LINE, long interspersed repeat element; LTR, long terminal repeat; SINE, short interspersed repeat element.

*Calculated using repeats identified with RepeatMasker (A. F. Smit, unpublished data; see <http://www.repeatmasker.org>). Some subfamilies cannot be distinguished using this method and were collapsed into their parent family.

†The number of plus signs indicates the approximate percentage of element insertions that exhibit RIP-like mutations (transition/transversion ratio >10): +++, >50%; ++, >10%; +, <10%; and —, no evidence of RIP.

‡Elements with different subclasses.

§The presence of truncated copies made it difficult to distinguish between subclasses.

||New elements identified in this study.

¶Previously reported sequence⁵⁰.

However, within the genome sequence there are numerous repetitive elements that show only transition mutations. This, combined with the presence of a gene in *M. grisea* predicted to encode a DNA methyltransferase homologous to the *RID* gene of *N. crassa*, which is required for RIP, indicates recent and possibly ongoing RIP⁴³. Second, the experimental evidence for RIP in *M. grisea* also demonstrated that the process is both less efficient at recognizing repeats and less severe in the number of mutations induced than in *N. crassa*⁴². Therefore, it is possible that many transposable elements in *M. grisea* simply escape damage by RIP. Finally, because RIP is only known to operate during the sexual cycle, the lack of RIP-like mutations in most repetitive elements may reflect the predominance of vegetative propagation during the recent evolution of rice pathogenic strains of *M. grisea*.

Discussion

This is the first analysis of the genome of a plant pathogenic fungus. Our analysis of the *M. grisea* genome has allowed a much greater appreciation of the likely attributes required by a plant pathogenic fungus to invade and colonize a living host plant. *M. grisea* has a considerably diverse set of proteins involved in extracellular perception and signal transduction, an extensive array of secreted proteins and secondary metabolites, specifically adapted regulatory pathways controlling infection-related development, and a genome capable of generating considerable genetic variation even in the absence of sexual reproduction. New opportunities for disease control and novel targets—such as unique families of CFEM-GPCR surface receptors and secreted cysteine-rich proteins—for development of durable fungicides are already apparent from our analyses. Acquisition of the genome sequence of *M. grisea* and that of its host, rice^{44,45}, coupled with the genetic and experimental tractability of this pathosystem, will enable a systems biology approach to the study of a plant–fungal interaction. Large-scale, insertional mutagenesis projects, transcriptional profiling and proteomic analysis are already in progress and offer the possibility of an enhanced understanding of the processes by which a fungus causes plant disease.

The genome sequence of *M. grisea* illustrates an extraordinary feature of the fungal kingdom, namely the extent of sequence diversity between what are thought of as closely related species from a taxonomic perspective. For example, although *M. grisea* and *N. crassa* are both pyrenomycetes and are often studied using similar methods in the same laboratory, they exhibit a degree of sequence diversity similar to that found between human and *Xenopus*⁴⁶. □

Methods

Strain and growth conditions

Rice-infecting *M. grisea* strain 70-15 (ref. 47) (Fungal Genetics Stock Center 8958) was grown in liquid complete medium, and DNA was extracted as previously described⁴⁸.

Sequencing and assembly

Plasmid (4-kb inserts) and fosmid (40-kb inserts) libraries were generated and end-sequenced as described (<http://www.broad.mit.edu/annotation/fungi/magnaporthe/assembly.html#clones>). Bacterial artificial chromosome (BAC) libraries were generated and end-sequenced as described⁴⁹. The draft genome sequence was assembled using Arachne (<http://www.broad.mit.edu/wga/>). Mapped genetic markers were associated with sequence scaffolds through hybridization to end-sequenced BACs (Supplementary Methods S1). ESTs were aligned to the genome using a BLAST *e*-value cutoff of $\leq 10^{-20}$. Assembly version 2 was used for all subsequent analyses.

Annotation and analysis

Automated gene predictions and annotations were performed using Calhoun (http://www.broad.mit.edu/annotation/fungi/magnaporthe/gene_finding.html). Gene predictions were performed using FGENESH/FGENESH+ trained on *M. grisea* sequences (SoftBerry) and GENES (Sanger Center) and validated against 65 characterized *M. grisea* genes. Additional information and gene identification numbers for genes featured in this article are presented in Supplementary Table S8.

To perform co-linearity analyses, amino acid identity between *M. grisea* and *N. crassa* was first determined by comparing the predicted proteins from each fungus using BLASTP. Homologues with the best hit were aligned using ClustalW and the amino acid per cent identity for each pair was calculated. *M. grisea* and *N. crassa* genes were considered

to belong to a conserved cluster if there was less than 10 kb between any two genes in the cluster. Homologous genes in clusters between species were accepted if alignments spanned $\geq 60\%$ of both genes and the alignment score was within 80% of the top score for either of the pair of genes. In this analysis, a gene may be placed in more than one cluster. No attempt was made to identify or resolve these cases.

Blastclust (<http://ftp.ncbi.nlm.nih.gov/blast/executables/LATEST/>) was used to cluster predicted peptide sequences from *M. grisea*, *N. crassa* and *A. nidulans* into families using threshold limits of 30% identity and 80% length overlap. Functional annotations were manually curated by examining blast searches to the GenBank NR protein database, SwissProt and a database of *M. grisea* repetitive sequences. A heterogeneity G-test was performed to identify families with significant content differences between species. Pair-wise post-hoc tests were performed using the Bonferroni correction for multiple comparisons. Families with significant similarity to proteins encoded by transposable elements were removed.

To identify G-protein-coupled receptor-like proteins, known GPCR sequences, including ones present in GPCRDB (www.gpcr.org/7tm/), were blasted (e -value limit of $\leq 10^{-9}$) against the *M. grisea* predicted proteins. These, in addition to candidates identified from an Interpro scan, were confirmed to have seven transmembrane spans by TMPRED (http://www.ch.embnet.org/software/TMPRED_form.html), Phobius (<http://phobius.cgb.ki.se/>) and TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>). Default settings were used. *N. crassa* and *A. nidulans* CFEM-containing GPCRs were identified by BLASTP.

Putative polyketide synthases and non-ribosomal peptide synthetases were identified by comparison of the genome sequence to known sequences in GenBank from *A. nidulans*, *Leptosphaeria maculans*, *Nectria haematococca*, *Gibberella moniliformis* and *Cochliobolus heterostrophus*.

Microarray experiments were performed using conidia germinated in water on either the hydrophobic or hydrophilic side of GelBond. At 7 and 12 h after germination, samples were flash frozen with liquid nitrogen, scraped from the support, ground and RNA extracted using Sigma Trizol reagent. RNA was similarly extracted from ungerminated conidia. RNA from two biological replications of each treatment was pooled in equal amounts and labelled with Cy3 and Cy5 dyes using Agilent Technologies low input linear amplification kit. Samples were hybridized to the Agilent *M. grisea* whole genome oligo array (product G4137A) using manufacturer protocols and reagents. Hybridizations were performed in an interlaced loop where each treatment was paired with every other. A total of ten hybridizations were performed, each treatment was used in four hybridizations (two with Cy3 and two with Cy5). Spot fluorescence was normalized using Lowess within and between slides, and gene expression profiles were analysed in GeneSpring. Microarray data may be accessed through NCBI GEO (<http://www.ncbi.nlm.nih.gov/projects/geo/>) accession GSE1945.

Additional information concerning genome sequencing and analysis can be found at <http://www.broad.mit.edu/annotation/fungi/magnaporthe/>.

Received 4 November 2004; accepted 7 February 2005; doi:10.1038/nature03449.

- Valent, B. & Chumley, F. G. Molecular genetic analysis of the rice blast fungus *Magnaporthe grisea*. *Annu. Rev. Phytopathol.* **29**, 443–467 (1991).
- Talbot, N. J. On the trail of a cereal killer: investigating the biology of *Magnaporthe grisea*. *Annu. Rev. Microbiol.* **57**, 177–202 (2003).
- Zeigler, R. S., Leong, S. A. & Teeng, P. S. *Rice Blast Disease* (CAB International, Wallingford, 1994).
- Hamer, J. E., Howard, R. J., Chumley, F. G. & Valent, B. A mechanism for surface attachment in spores of a plant pathogenic fungus. *Science* **239**, 288–290 (1988).
- Dean, R. A. Signal pathways and appressorium morphogenesis. *Annu. Rev. Phytopathol.* **35**, 211–234 (1997).
- de Jong, J. C., McCormack, B. J., Smirnov, N. & Talbot, N. J. Glycerol generates turgor in rice blast. *Nature* **389**, 244–245 (1997).
- Sesma, A. & Osbourn, A. E. The rice blast pathogen undergoes developmental processes typical of root-infecting fungi. *Nature* **431**, 582–586 (2004).
- Sweigard, J. A. *et al.* in *Genetic Maps* (ed. O'Brien, S. J.) 3.112 (Cold Spring Harbor, New York, 1993).
- Nitta, N., Farman, M. L. & Leong, S. A. Genome organization of *Magnaporthe grisea*: integration of genetic maps, clustering of transposable elements and identification of genome duplications and rearrangements. *Theor. Appl. Genet.* **95**, 20–32 (1997).
- Takano, Y., Choi, W. B., Mitchell, T. K., Okuno, T. & Dean, R. A. Large scale parallel analysis of gene expression during infection-related morphogenesis of *Magnaporthe grisea*. *Mol. Plant Pathol.* **4**, 337–346 (2003).
- Ebbold, D. J. *et al.* Gene discovery and gene expression in the rice blast fungus, *Magnaporthe grisea*: Analysis of expressed sequence tags. *Mol. Plant Microbe Interact.* **17**, 1337–1347 (2004).
- Galagan, J. E. *et al.* The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* **422**, 859–868 (2003).
- Idnurm, A. & Howlett, B. J. Pathogenicity genes of phytopathogenic fungi. *Mol. Plant Pathol.* **2**, 241–255 (2001).
- Schmidt, R. Synteny: recent advances and future prospects. *Curr. Opin. Plant Biol.* **3**, 97–102 (2000).
- Fredriksson, R., Lagerstrom, M. C., Lundin, L. G. & Schiöth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **63**, 1256–1272 (2003).
- Bolker, M. Sex and crime: heterotrimeric G proteins in fungal mating and pathogenesis. *Fungal Genet. Biol.* **25**, 143–156 (1998).
- Kulkarni, R. D., Kelkar, H. S. & Dean, R. A. An eight-cysteine-containing CFEM domain unique to a group of fungal membrane proteins. *Trends Biochem. Sci.* **28**, 118–121 (2003).
- DeZwaan, T. M., Carroll, A. M., Valent, B. & Sweigard, J. A. *Magnaporthe grisea* Pth11p is a novel plasma membrane protein that mediates appressorium differentiation in response to inductive surface cues. *Plant Cell* **11**, 2012–2030 (1999).
- Xu, J. R. & Hamer, J. E. MAP kinase and cAMP signalling regulate infection structure formation and pathogenic growth in the rice blast fungus *Magnaporthe grisea*. *Genes Dev.* **10**, 2696–2706 (1996).
- Xu, J. R. MAP kinases in fungal pathogens. *Fungal Genet. Biol.* **31**, 137–152 (2000).
- Ptashe, M. & Gann, A. Imposing specificity on kinases. *Science* **299**, 1025–1027 (2003).
- Mitchell, T. K. & Dean, R. A. The cAMP-dependent protein kinase catalytic subunit is required for appressorium formation and pathogenesis by the rice blast pathogen *Magnaporthe grisea*. *Plant Cell* **7**, 1869–1878 (1995).
- Theiringer, R. & Kunau, W. H. The β -oxidation system in catalase-free microbodies of the filamentous fungus *Neurospora crassa*. Purification of a multifunctional protein possessing 2-enoyl-CoA hydratase, L-3-hydroxyacyl-CoA dehydrogenase, and 3-hydroxyacyl-CoA epimerase activities. *J. Biol. Chem.* **266**, 13110–13117 (1991).
- Thines, E., Weber, R. W. S. & Talbot, N. J. MAP kinase and protein kinase A-dependent mobilization of triacylglycerol and glycogen during appressorium turgor generation by *Magnaporthe grisea*. *Plant Cell* **12**, 1703–1718 (2000).
- Sweigard, J. A., Chumley, F. G. & Valent, B. Disruption of a *Magnaporthe grisea* cutinase gene. *Mol. Genet.* **232**, 183–190 (1992).
- Van den Berg, H. A. *et al.* Natural disulfide bond-disrupted mutants of AVR4 of the tomato pathogen *Cladosporium fulvum* are sensitive to proteolysis, circumvent Cf-4-mediated resistance, but retain their chitin-binding ability. *J. Biol. Chem.* **278**, 27340–27346 (2003).
- Nimchuk, Z., Eulgem, T., Holt, B. F. III & Dangl, J. L. Recognition and response in the plant immune system. *Annu. Rev. Genet.* **37**, 579–609 (2003).
- Jia, Y., McAdams, S. A., Bryan, G. T., Hershey, H. P. & Valent, B. Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *EMBO J.* **19**, 4004–4014 (2000).
- Fellbrich, G. *et al.* NPP1, a Phytophthora-associated trigger of plant defense in parsley and *Arabidopsis*. *Plant J.* **32**, 375–390 (2002).
- Joosten, M. H. A. J., Cozijnsen, T. J. & de Wit, P. J. G. M. Host resistance to the fungal tomato pathogen lost by a single base-pair change in the avirulence gene. *Nature* **367**, 384–386 (1994).
- Moss, J., Stanley, S. J., Vaughan, M. & Tsuji, T. Interaction of ADP-ribosylation factor with *Escherichia coli* enterotoxin that contains an inactivating lysine 112 substitution. *J. Biol. Chem.* **268**, 6383–6387 (1993).
- Steiner-Lange, S. *et al.* Differential defense reactions in leaf tissues of barley in response to infection by *Rhynchosporium secalis* and to treatment with a fungal avirulence gene product. *Mol. Plant Microbe Interact.* **16**, 893–902 (2003).
- Wolpert, T. J., Dunkle, L. D. & Ciuffetti, L. M. Host-selective toxins and avirulence determinants: what's in a name? *Annu. Rev. Phytopathol.* **40**, 251–285 (2002).
- Kroken, S., Glass, N. L., Taylor, J. W., Yoder, O. C. & Turgeon, B. G. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc. Natl Acad. Sci. USA* **100**, 15670–15675 (2003).
- Bohner, H. U. *et al.* A putative polyketide synthase/peptide synthetase from *Magnaporthe grisea* signals pathogen attack to resistant rice. *Plant Cell* **16**, 2499–2513 (2004).
- Bonman, J. M. Durable resistance to rice blast disease—environmental influences. *Euphytica* **63**, 115–123 (1992).
- Farman, M. L. *et al.* Analysis of the structure of the AVR1-CO39 avirulence locus in virulent rice-infecting isolates of *Magnaporthe grisea*. *Mol. Plant Microbe Interact.* **15**, 6–16 (2002).
- Kang, S., Lebrun, M.-H., Farrall, L. & Valent, B. Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. *Mol. Plant Microbe Interact.* **14**, 671–674 (2001).
- Thon, M. R., Martin, S. L., Goff, S., Wing, R. A. & Dean, R. A. BAC end sequences and a physical map reveal transposable element content and clustering patterns in the genome of *Magnaporthe grisea*. *Fungal Genet. Biol.* **41**, 657–666 (2004).
- Farman, M. L. Meiotic deletion at the *BUF1* locus of the fungus *Magnaporthe grisea* is controlled by interaction with the homologous chromosome. *Genetics* **160**, 137–148 (2002).
- Kang, S., Sweigard, J. A. & Valent, B. The PWL host specificity gene family in the rice blast fungus, *Magnaporthe grisea*. *Mol. Plant Microbe Interact.* **8**, 939–948 (1995).
- Iked, K. *et al.* Repeat-induced point mutation (RIP) in *Magnaporthe grisea*: implications for its sexual cycle in the natural field context. *Mol. Microbiol.* **45**, 1355–1364 (2002).
- Freitag, M., Williams, R. L., Kothe, G. O. & Selker, E. U. A cytosine methyltransferase homologue is essential for repeat-induced point mutation in *Neurospora crassa*. *Proc. Natl Acad. Sci. USA* **99**, 8802–8807 (2002).
- Yu, J. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**, 79–92 (2002).
- Goff, S. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92–100 (2002).
- Hedges, S. B. The origin and evolution of model organisms. *Nature* **3**, 838–849 (2002).
- Chao, C. & Ellingboe, A. Selection for mating competence in *Magnaporthe grisea* pathogenic to rice. *Can. J. Bot.* **69**, 2130–2134 (1991).
- Weiland, J. J. Rapid procedure for extraction of DNA from fungal spores and mycelium. *Fungal Genet. Newslett.* **44**, 60–63 (1997).
- Zhu, H., Blackmon, B. P., Sasinowski, M. & Dean, R. A. Physical map and organization of chromosome 7 in the rice blast fungus, *Magnaporthe grisea*. *Genome Res.* **9**, 739–750 (1999).
- Nakayashiki, H. *et al.* Pyret, a Ty3/Gypsy retrotransposon in *Magnaporthe grisea* contains an extra domain between the nucleocapsid and protease domains. *Nucleic Acids Res.* **29**, 4106–4113 (2001).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements The authors acknowledge the USDA-CSREES and the National Science Foundation for funding this work. We thank all the members of the Dean Laboratory at NCSU, and members of the Broad Institute Sequencing Platform, Assembly and Annotation teams, and members at each collaborating laboratory. We also thank the rice blast research community at large. We acknowledge other fungal research communities, in particular the *A. nidulans* community, for making it possible to have access to genome sequence information before publication.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.A.D. (Ralph_Dean@ncsu.edu). The whole-genome shotgun data has been deposited at GenBank/EMBL/DDJB under the project accession number AACU00000000.

Molecular mechanisms of kinetochore capture by spindle microtubules

Kozo Tanaka¹, Naomi Mukae¹, Hilary Dewar¹, Mark van Breugel², Euan K. James¹, Alan R. Prescott¹, Claude Antony³ & Tomoyuki U. Tanaka¹

¹School of Life Sciences, University of Dundee, Wellcome Trust Biocentre, Dundee DD1 5EH, UK

²Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany

³European Molecular Biology Laboratory, D-69117 Heidelberg, Germany

For high-fidelity chromosome segregation, kinetochores must be properly captured by spindle microtubules, but the mechanisms underlying initial kinetochore capture have remained elusive. Here we visualized individual kinetochore–microtubule interactions in *Saccharomyces cerevisiae* by regulating the activity of a centromere. Kinetochores are captured by the side of microtubules extending from spindle poles, and are subsequently transported poleward along them. The microtubule extension from spindle poles requires microtubule plus-end-tracking proteins and the Ran GDP/GTP exchange factor. Distinct kinetochore components are used for kinetochore capture by microtubules and for ensuring subsequent sister kinetochore bi-orientation on the spindle. Kar3, a kinesin-14 family member, is one of the regulators that promote transport of captured kinetochores along microtubules. During such transport, kinetochores ensure that they do not slide off their associated microtubules by facilitating the conversion of microtubule dynamics from shrinkage to growth at the plus ends. This conversion is promoted by the transport of Stu2 from the captured kinetochores to the plus ends of microtubules.

Sister chromatid segregation to opposite poles of the cell during mitosis is crucial for maintenance of genetic integrity in eukaryotic cells. For high-fidelity chromosome segregation, kinetochores must be properly captured by spindle microtubules¹. In vertebrate cells, after nuclear envelope breakdown, kinetochores are initially captured by the lattice (the lateral surface, rather than the tip) of a single microtubule extending from a spindle pole^{2–4}. Kinetochores are then transported poleward along the surface of a microtubule that terminates distal to the kinetochores. During their poleward travel, kinetochores interact with more microtubules, the plus ends of which subsequently become embedded into the kinetochore plates (end-on attachment). Eventually, each sister kinetochore attaches to the plus ends of microtubules from opposite spindle poles. In this way, sister kinetochores bi-orient on the mitotic spindle before anaphase onset. Because kinetochore capture and transport have only been visualized in very few cell types and on limited occasions, mechanisms of kinetochore capture by microtubules have remained elusive.

Recently, centromere movements were visualized in the budding yeast *S. cerevisiae* by marking individual centromeres with green fluorescent protein (GFP)^{5–8}. When sister kinetochores bi-orient, the centromere GFP signals split on the pre-anaphase spindle while sister chromatids are still being held together by cohesion along their arms. However, it is poorly understood how kinetochores interact with microtubules before they bi-orient on the spindle. In *S. cerevisiae*, centromeres are tethered by microtubules to spindle pole bodies (SPBs) during most of the cell cycle^{9,10}. Nonetheless, centromeres are released from, and recaptured by, microtubules during a brief period in S phase, probably due to kinetochore disassembly and reassembly upon centromere DNA replication^{10,11}. Kinetochore capture by microtubules can therefore be studied in budding yeast. However, because all kinetochores are captured by microtubules during a short time period and within a small space around SPBs, it has been difficult to resolve individual kinetochore capture events using light microscopy.

Visualizing kinetochore capture and transport

To visualize individual kinetochore–microtubule interactions in *S. cerevisiae*, we devised the following system (Fig. 1a). We replaced *CEN3* on chromosome III with *CEN3* under control of the *GAL1-10*

promoter¹², to conditionally inactivate and activate the centromere by turning on and off transcription from the *GAL1-10* promoter in the presence of galactose and glucose, respectively. In addition, to enable detection of *CEN3* in live cells, we marked the *CEN3* sequence by the adjacent insertion of a *tet* operator array that is bound by Tet repressors fused with GFP¹³. Furthermore, microtubules were visualized by expressing α -tubulin (*TUB1*) fused with yellow fluorescent protein (YFP). To conditionally arrest cells in metaphase, the *CDC20* gene, which is required for sister chromatid separation and for anaphase onset¹⁴, was placed under control of the *MET3* promoter, which can be turned off in the presence of methionine. To observe individual kinetochore–microtubule interactions, we wanted to detect the GFP-marked *CEN3* distant from the spindle and from all other centromeres locating on the spindle. To this end, we inactivated the *CEN3* upon the release of cells from α -factor arrest. Simultaneously, we also depleted Cdc20 protein levels to arrest cells in metaphase. Once cells were arrested in metaphase, we reactivated *CEN3* by turning off the adjacent *GAL1-10* promoter while cells were still in metaphase, and followed the behaviour of *CEN3*.

Two hours after release from α -factor arrest, cells were arrested in metaphase and *CEN3* was distant from the spindle in 60–70% of cells. After *CEN3* reactivation, the GFP-marked centromere interacted with microtubules extending from a spindle pole, and moved poleward along the microtubules (Fig. 1b and Supplementary Video S1).

To study kinetochore capture and transport in detail, we wanted to observe longer microtubules because their dynamics can be characterized more precisely. We achieved this by arresting cells in metaphase for longer periods, which led to elongation of the nucleus between buds and mother cell bodies. This elongation was probably due to earlier back-and-forth motions of the spindle between the two cell bodies, and was not due to cells leaking into anaphase, because the spindle length stayed short. When *CEN3* and the spindle were located on opposite sides of such elongated nuclei, microtubules extended for a longer distance from a spindle pole, and captured *CEN3* after *CEN3* reactivation (see Fig. 1c, Supplementary Video S2 and Supplementary Fig. S1). *CEN3* seemed to be captured laterally by microtubules and transported poleward along them. Sometimes *CEN3* motion halted or occurred

anti-poleward for a short period, but eventually poleward motion predominated (for example, Fig. 3a, bottom). Shortly after *CEN3* reached a spindle pole (<5 min), the majority of its GFP signal split on the spindle, indicating that sister *CEN3*s had bi-oriented on the spindle^{5–8}. During transport along microtubules, *CEN3* sometimes transiently detached from the microtubules, or changed its associated microtubules to different ones extending from either the same or the opposite pole (Supplementary Fig. S2, Supplementary Video S3). (See Supplementary Note 1.)

To rule out the possibility that kinetochores are pulled by end-on attached microtubules that might have overlapped longer microtubules, we first visualized Bik1 and Bim1, both of which localized to the plus ends of microtubules extending from spindle poles. As *CEN3* moved poleward along microtubules, Bik1 and Bim1 signals

always localized distal to *CEN3*, but not at *CEN3* (Supplementary Fig. S3 and Supplementary Note 2). Second, in many cases, the intensity of YFP–Tub1 signals was almost constant along the length of the *CEN3*-associated microtubule, from the spindle pole to the plus end (Supplementary Fig. S4a and Supplementary Video S4). These observations confirm that *CEN3* is indeed captured and transported along the sides of microtubules.

To address whether *CEN3* is captured by a single microtubule, we compared YFP–Tub1 signals from *CEN3*-associated microtubules and from the thinnest cytoplasmic microtubules (Supplementary Fig. S4b), which constitute the majority of cytoplasmic microtubules and are thought to be singular^{15,16}. For the majority of *CEN3*-associated microtubules, the intensity of YFP–Tub1 signals was similar to that of the thinnest cytoplasmic microtubules. Thus, many kinetochore-associated microtubules are probably singular, although sometimes other microtubules grew from a spindle pole, overlapping the microtubule that had extended earlier. (See Supplementary Note 3.)

We then addressed whether *CEN3* facilitates nuclear microtubule extension from spindle poles in its direction before capture, or whether microtubule extension occurs in random directions within the nucleus. To distinguish nuclear microtubules from cytoplasmic microtubules, we labelled the nuclear periphery and used a marker for cytoplasmic microtubules (Supplementary Fig. S5). We compared the number of nuclear microtubules that extended towards *CEN3* and away from it (Supplementary Fig. S6). The direction of nuclear microtubules, initially extending from spindle poles, was largely unaffected by the position of free kinetochores that had not yet been captured by microtubules. (See Supplementary Note 4.)

Nuclear microtubule extension and kinetochore capture

We studied the roles of candidate regulators in generating nuclear microtubules from spindle poles and in kinetochore capture by these microtubules. In *S. cerevisiae*, kinetochores contain a number of protein complexes¹⁷: the CBF3 complex (containing Ndc10) directly binds the centromere DNA; the Dam1 complex (containing Spc34 and Ask1) is thought to provide a direct interface for microtubule attachments to kinetochores in metaphase; the Ndc80 (containing Spc24 and Spc25), Mtw1 (containing Dsn1) and Ctf19 (containing Okp1) complexes are probable components bridging a gap between the CBF3 complex with centromeres and the Dam1 complex with microtubules; and finally, the Ipl1–Sli15 complex (orthologue of the metazoan Aurora B–INCENP complex) ensures sister kinetochore bi-orientation^{10,18}. Kinetochore–microtubule interactions might also be regulated by the microtubule plus-end tracking proteins (+TIPs) Bim1, Bik1, Stu1 and Stu2 (orthologues of vertebrate EB1, CLIP170, CLASP and XMAP215/ch-TOG, respectively)¹⁹, and by the small GTPase Ran and its regulators Prp20 (also called Mtr1) and Rna1, which operate as a nuclear Ran GDP/GTP exchange factor (GEF) and a cytoplasmic Ran GTPase-activating protein (GAP), respectively²⁰. Although Ran GEF and Ran GAP regulate nuclear import and export in yeast and metazoan cells, they also regulate spindle formation and kinetochore function after nuclear envelope breakdown (open mitosis) in metazoan cells^{21,22}. It is not clear whether this is the case in budding yeast, in which the nuclear envelope remains intact throughout the cell cycle (closed mitosis)⁹.

We first determined which factors facilitate nuclear microtubule extension from spindle poles. We counted the number of nuclear microtubules in elongated nuclei of mutants of the above regulators. *ndc10-1*, *spc24-1*, *spc25-1*, *dsn1-7*, *mtw1-11*, *okp1-5*, *spc34-3*, *ask1-3*, *dam1-1*, *ipl1-321*, *sli15-3* and *rna1-1* mutants showed similar numbers of nuclear microtubules compared to ‘wild-type’ cells (cells without mutation; Fig. 2a and data not shown). In contrast, the number of nuclear microtubules was significantly reduced in *prp20-1*, *mtr1-1*, *stu2-10*, *bim1Δ* and *bik1Δ* mutants (Fig. 2a, b and

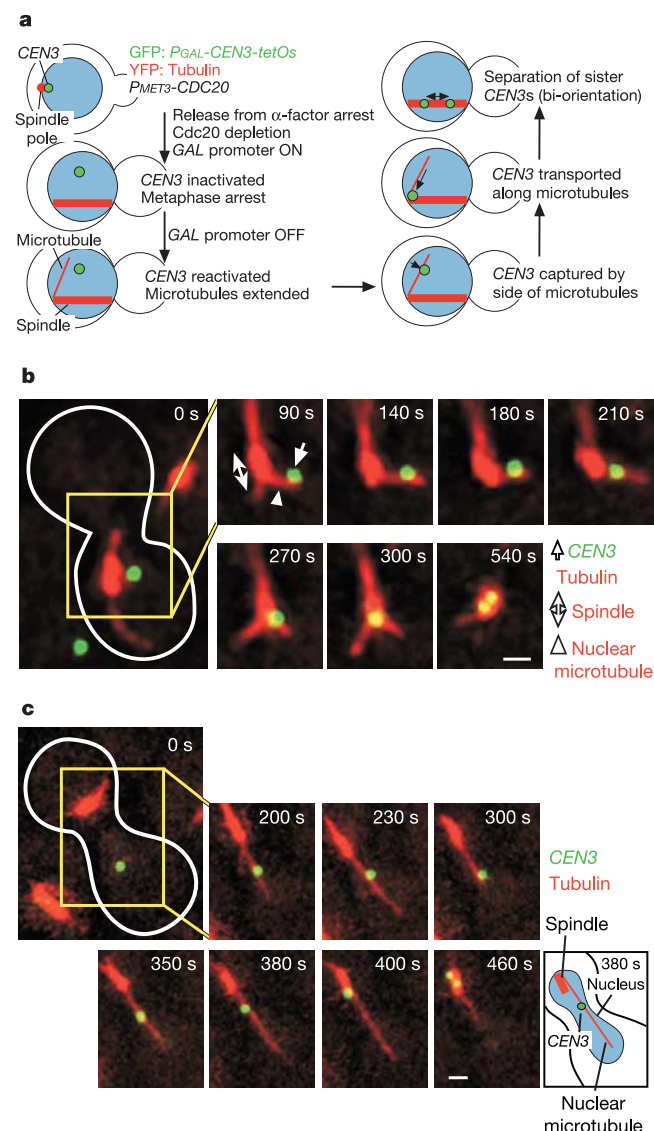


Figure 1 Visualizing kinetochore capture and transport along microtubules.

a, Experimental system and diagrams showing kinetochore capture. **b**, **c**, Time-lapse images. *P_{MET3}-CDC20 P_{GAL}-CEN3-tetOs TetR-GFP YFP-TUB1* cells (T3531) were treated with α -factor in methionine drop-out medium with raffinose for 2.5 h, and then released to YP medium containing galactose, raffinose and 2 mM methionine. After 2 h (**b**) or 3 h (**c**), cells were suspended in synthetic complete medium containing glucose and methionine. Images were collected every 10 s using separate channels for GFP (*CEN3*; green) and YFP (tubulin; red). Zero time is set arbitrarily for the first panel, in which the cell shape is outlined in white. Scale bar, 1 μ m. See Supplementary Videos S1 and S2.

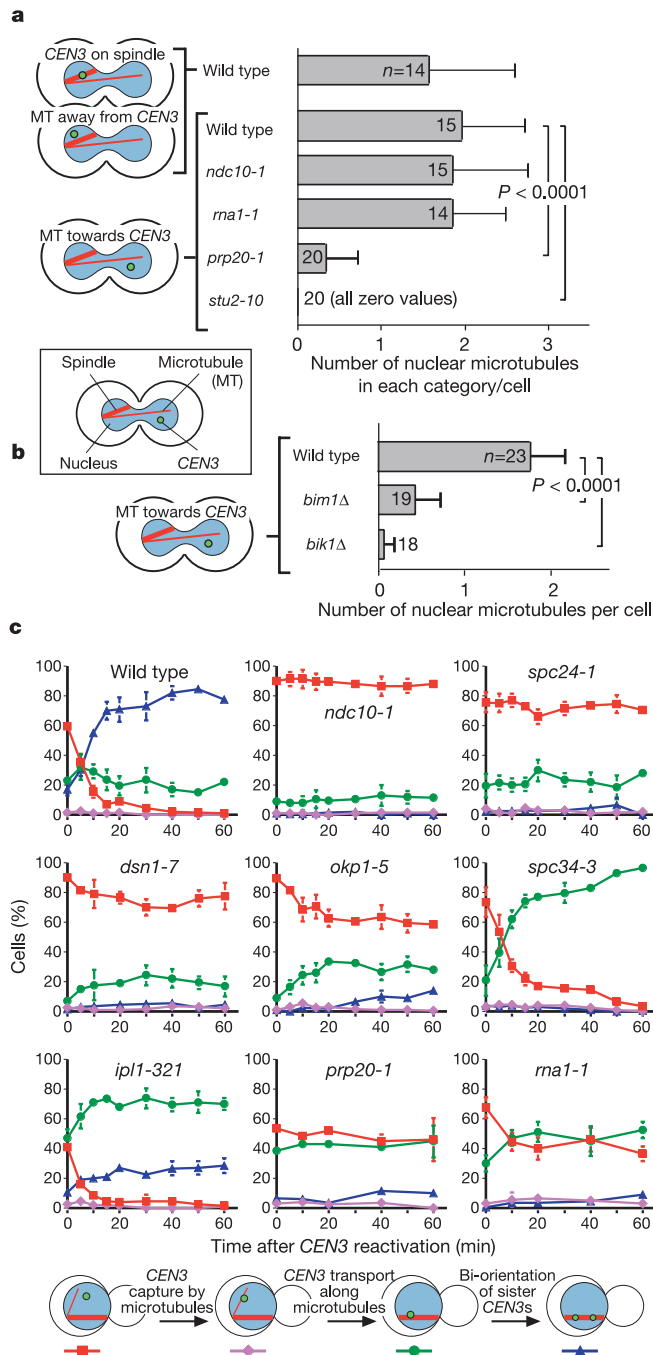


Figure 2 Mechanisms for nuclear microtubule extension and for kinetochore capture. **a, b**, Number of nuclear microtubules. *P_{MET3}-CDC20 P_{GAL}-CEN3-tetO TetR-GFP YFP-TUB1 YFP-NIC96 KIP2-4GFP* (T3270 or 'wild type') cells, and cells with the same genotype but including different mutations (*ndc10-1*, *rna1-1*, *prp20-1*, *stu2-10*, *bim1Δ* or *bik1Δ*) were treated as in Fig. 1c (**b**), or as in Fig. 1c but with cultures shifted to 35 °C (restrictive temperature) 30 min before transfer to glucose-containing medium (**a**). Time-lapse images were collected at 35 °C (**a**) or at 23 °C (**b**) every 10 s for 30 min. Only microtubules that extended from the spindle to the other side (towards the right in diagrams) of elongated nuclei were counted, after cells were classified by CEN3 location at the beginning of time-lapse. *n* represents the number of observed cells. **c**, Time course for CEN3 capture. T3531 'wild-type' cells (see Fig. 1 legend) and strains with the same genotype but including different mutations (*ndc10-1*, *spc24-1*, *dsn1-7*, *okp1-5*, *spc34-3*, *ipl1-321*, *prp20-1*, *ma1-1*) were treated as in **a** and fixed at the indicated time points (see Supplementary Note 17e). Coloured lines show the percentage of cells in which CEN3 has not yet been captured by microtubules (red), in which CEN3 is on the microtubule extending from a spindle pole (magenta), in which sister CEN3s are not separated (green) or separated (blue) on the spindle. See Supplementary Note 4a and Supplementary Fig. S5 for information about YFP-NIC96 KIP2-4GFP.

data not shown). Thus, Ran GEF and +TIPs (Stu2, Bim1 and Bik1) are required to facilitate nuclear microtubule extension from spindle poles, and mutations to kinetochore components do not greatly affect this process. Ran GEF probably facilitates nuclear microtubule extension by increasing the nuclear concentration of RanGTP, because a Ran mutant also showed less frequent microtubule extension (data not shown). RanGTP may directly regulate microtubule dynamics in closed mitosis, as it does in open mitosis in *Xenopus* egg extracts^{23,24}. At the very least, defects in nuclear microtubule extension in Ran GEF mutants were not secondary to

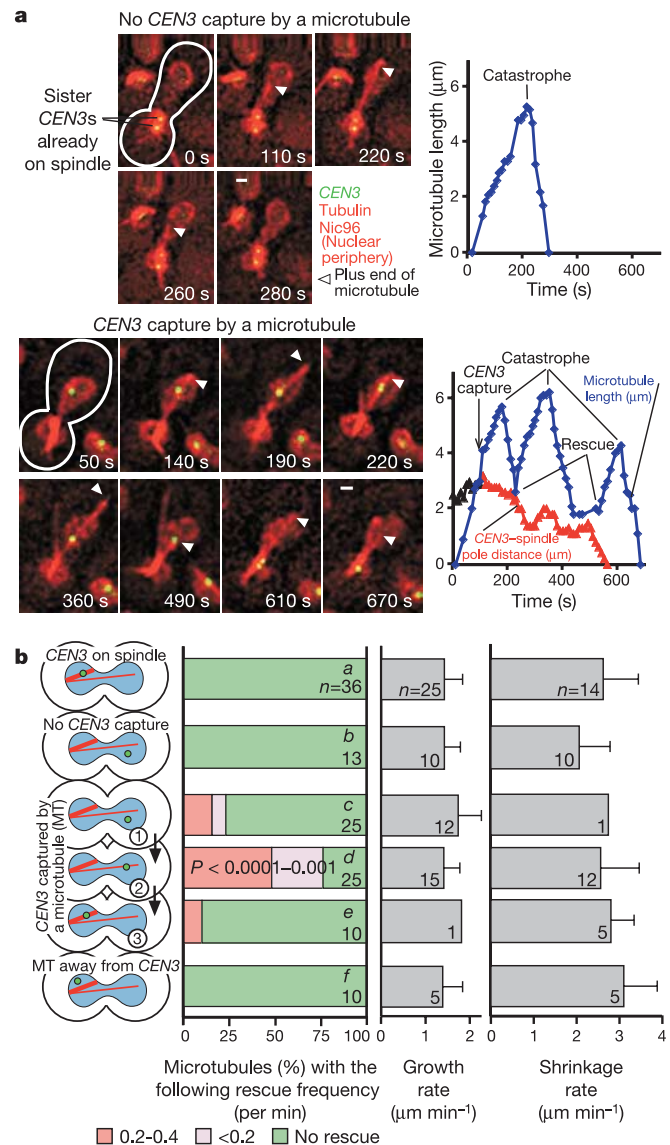


Figure 3 Captured kinetochores facilitate microtubule rescue. T3270 cells (see Fig. 2 legend) were treated as in Fig. 1c. **a**, Representative time-lapse images of nuclear microtubules that did not capture CEN3 (top) and did capture CEN3 (bottom). Scale bar, 1 μm. Graphs show microtubule length (blue). The lower graph also shows the distance of CEN3 from a spindle pole before (black) and after (red) CEN3 was captured by the microtubule. See Supplementary Videos S5 and S6. **b**, Dynamics of nuclear microtubules. Only microtubules that extended from the spindle towards the other side (towards the right in diagrams) of elongated nuclei were analysed, after cells were classified by CEN3 location at the beginning of time-lapse and the fate of CEN3 during observation. The rescue frequency of microtubules in bar *d* was significantly higher than that in *a*, *b*, *c* and *f* ($P < 0.0001$), and in *e* ($P < 0.001$). *n* represents the number of observed microtubules (left graph) or the number of observed events (middle and right graphs).

the nuclear import defects in these mutants. (See Supplementary Note 5.)

We then studied the kinetics of kinetochore capture by microtubules after *CEN3* reactivation in the above mutants. As expected, *CEN3* capture was delayed in yeast containing mutations to Ran GEF (Fig. 2c) and +TIPs (data not shown), where fewer nuclear microtubules extended. In *ndc10-1* (CBF3 complex), *spc24-1* and *spc25-1* (Ndc80 complex), *dsn1-7* and *mtw1-11* (Mtw1 complex), *okp1-5* (Ctf19 complex) and *rna1-1* (Ran GAP) mutants, the frequency of nuclear microtubule extension was comparable to wild-type cells, but a lack or significant delay in kinetochore capture was observed (Fig. 2c and data not shown). In contrast, *ipl1-321* and *sli15-3* mutants showed no significant delay in *CEN3* capture by microtubules. *spc34-3*, *ask1-3* and *dam1-1* (Dam1 complex) mutants also showed normal kinetics of *CEN3* capture, except for 10–15% of cells in which *CEN3* was not captured, perhaps owing to spindle abnormality (for example, spindle disruption). *ipl1-321*, *sli15-3*, *spc34-3*, *ask1-3* and *dam1-1* mutants showed no significant defects in *CEN3* transport along microtubules (data not shown). However, despite almost normal kinetics for *CEN3* capture and transport, these mutants showed a lack of or a severe delay in splitting the GFP signals of *CEN3* on the mitotic spindle after *CEN3* reached a spindle pole (Fig. 2c and data not shown).

Thus, the CBF3, Ndc80, Mtw1 and Ctf19 complexes, but not the Dam1 or Ipl1 complex, are necessary for kinetochore capture by the side of microtubules. In agreement with earlier reports^{10,25–29}, the Dam1 and Ipl1 complexes are required to ensure sister kinetochore bi-orientation. Ran GAP might facilitate kinetochore capture by microtubules, either by promoting nuclear import/export of unidentified regulators or by a more direct mechanism. (See Supplementary Note 6.)

Captured kinetochores facilitate microtubule rescue

After catastrophe (conversion from growth to shrinkage), nuclear microtubules shrank faster ($2\text{--}3\ \mu\text{m min}^{-1}$) than the average velocity of poleward kinetochore transport ($0.5\text{--}2.0\ \mu\text{m min}^{-1}$), whether or not the microtubules were associated with kinetochores (Fig. 3). One might therefore assume that the plus ends of shrinking microtubules would pass by the kinetochores, causing the kinetochores to slide off the microtubules. However, this never happened. To address the reason for it, we compared the dynamics of nuclear microtubules that did and did not capture *CEN3*. Microtubules did not show significant differences in growth rate or shrinkage rate while they were associated with *CEN3*. However, rescue (conversion from shrinkage to growth) was only observed for microtubules that were associated with *CEN3* at some stage, and not for other microtubules (Fig. 3 and Supplementary Videos S5 and S6). Owing to this microtubule rescue, the life span of microtubules that captured *CEN3* was significantly prolonged while they were associated with *CEN3* (Supplementary Fig. S7). Thus, kinetochores do not slide off microtubules because captured kinetochores facilitate microtubule rescue and extend microtubule life span. (See Supplementary Note 7.)

We next addressed whether *CEN3*-capturing microtubules are dynamic at their plus ends or at their minus ends (that is, at the ends distal or proximal to spindle poles, respectively). For this, we marked a region proximal to a spindle pole on 12 putative single microtubules by photobleaching the YFP-Tub1 signal (Fig. 4a and Supplementary Video S7). In all cases, while *CEN3*-associated microtubules shrank and grew, or while *CEN3* moved along the microtubules, there was no appreciable change in distance between the marked region and the spindle pole. Thus, microtubules shrink and grow mainly at their plus ends while they are associated with kinetochores. This property is reminiscent of cytoplasmic and anaphase nuclear microtubules in budding yeast³⁰. The results also indicate that kinetochore transport is not due to microtubule flux towards a spindle pole¹.

Mechanisms of kinetochore-dependent microtubule rescue

How do kinetochores facilitate rescue of their associated microtubules? This rescue occurs preferentially when the microtubule plus ends come close to or reach the position of *CEN3* during microtubule shrinkage (Supplementary Fig. S8 and Supplementary Note 8). Nonetheless, microtubule rescue often happened even when *CEN3* was $1\ \mu\text{m}$ or more away from the microtubule plus

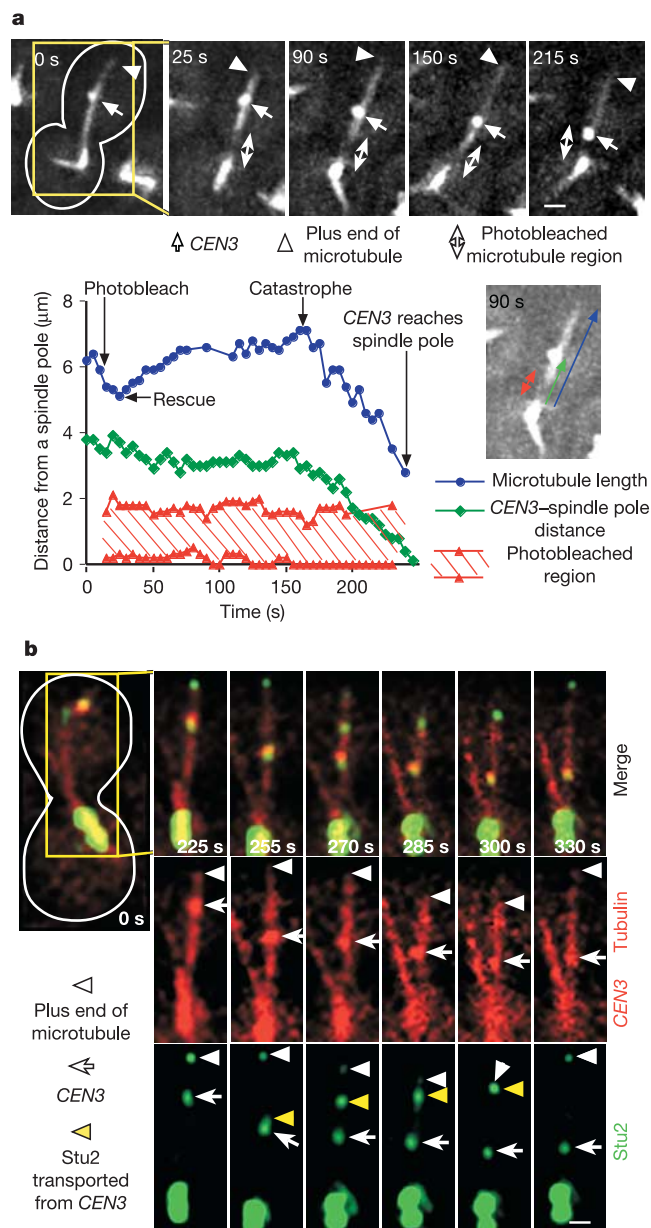


Figure 4 Microtubule rescue at the plus end coincides with the arrival of Stu2 protein transported from captured kinetochores. **a**, A kinetochore-associated microtubule shrinks and grows at its plus end. T3531 (see Fig. 1 legend) cells were treated as in Fig. 1c, except that GFP and YFP signals were collected together every 5 s. A microtubule region, proximal to a spindle pole, was photobleached between the 10 and 15 s time points while the microtubule was shrinking. See also Supplementary Video S7. **b**, Behaviour of Stu2 on a kinetochore-associated nuclear microtubule. *STU2-4GFP P_{MET3}-CDC20 P_{GAL}-CEN3-tetOs TetR-3CFP CFP-TUB1* cells (T3680) were treated as in Fig. 1c, except that cyan fluorescent protein CFP (*CEN3* and tubulin, red) and GFP (Stu2, green) signals were collected separately every 15 s. Yellow arrowheads show Stu2-4GFP signals moving from *CEN3* to the microtubule plus end, along the microtubule. See also Supplementary Fig. S9. Scale bar (**a**, **b**) $1\ \mu\text{m}$.

end (Supplementary Fig. S8). The distribution of rescue positions can be explained, for instance, if kinetochores hold regulators that facilitate microtubule rescue, and if these regulators are occasionally translocated from captured kinetochores to the microtubule plus ends.

What could be the identity of such regulators? Stu2, Bim1 and Bik1 are good candidates, because they affect microtubule dynamics at the plus end^{27,31–34} and facilitate the initial extension of nuclear microtubules (see above). Among the three candidates, only Stu2 (fused with 4 tandem copies of GFP) was detected on *CEN3* that had not yet been captured by microtubules (data not shown). Stu2 proteins localized to spindle poles and at the plus ends of extending nuclear microtubules (Fig. 4b). However, as microtubules reached their maximum length and subsequently shrank, Stu2 signals at the microtubule plus end gradually diminished (Fig. 4b, 225–285 s). Notably, after *CEN3* had been captured by the microtubule lattice, Stu2 was intermittently transported from *CEN3* along the microtubule towards the plus end (yellow arrowheads). The arrival of Stu2 at the microtubule plus end coincided with microtubule rescue (300–330 s). The transported Stu2 signal was not at the tip of another microtubule that might have overlapped with the existing longer microtubule (Supplementary Note 9b). We observed 24 rescues that occurred distal to *CEN3*, and all such rescues happened upon arrival of Stu2 (originating from *CEN3*) at the microtubule plus end (Supplementary Fig. S9). Arrival of Stu2 from *CEN3*, which occurred during microtubule shrinkage, almost always (in 24 out of 25 observations) led to immediate rescue (within 15 s) of the microtubule. Therefore, we propose that Stu2 is one of the mediators of kinetochore-dependent microtubule rescue. (See Supplementary Note 9.)

Kar3 kinesin is involved in kinetochore transport

What factors regulate kinetochore transport along the microtubule lattice? ATP-driven motor proteins could be involved. The *S. cerevisiae* genome encodes six kinesin family proteins (Cin8, Kar3, Kip1, Kip2, Kip3 and Smy1) and a single dynein heavy chain, Dyn1 (ref. 35). Among these motor proteins, Kar3 (a kinesin-14 family member) is thought to be the only nuclear microtubule minus-end directed motor. Kinetochore transport was studied in mutants containing single deletions of each of these motor proteins. Except for *kar3* deletion, none of these mutations significantly altered kinetochore transport (data not shown). A *cin8-3 kip1Δ* double mutant³⁶ (Cin8 and Kip1 belong to the same kinesin-5 family) also had no delay in kinetochore transport (data not shown). In the majority of *kar3Δ* cells, *CEN3* was transported poleward as in wild-type cells (Fig. 5b). However, in a small number of *kar3Δ* cells, *CEN3* was at a 'standstill', staying longer (≥ 21 min) at almost the same position on microtubules, with no appreciable movement. We further examined the dominant-negative *kar3-1* mutant (known as a 'rigour' mutant) that can bind microtubules but does not have motor activity, owing to an ATP hydrolysis defect^{37,38}. *kar3-1* cells showed longer and more frequent standstill than *KAR3*⁺ cells (Fig. 5, Supplementary Fig. S10a and Supplementary Videos S8 and S9). Moreover, overexpression of Kar3 made *CEN3* move poleward along microtubules, with shorter pauses compared with control cells (*KAR3*⁺) containing normal levels of Kar3 (Fig. 5, Supplementary Fig. S10b, Supplementary Videos S8 and S10). (See Supplementary Note 10.)

If Kar3 is involved in kinetochore transport, we would expect to find Kar3 loaded on kinetochores. Indeed, we found Kar3 association with centromeres using a chromatin immunoprecipitation assay. This association was (1) limited to a small centromeric region, (2) largely constant throughout the cell cycle, (3) dependent on Ndc10 and (4) probably independent of microtubule attachment to kinetochores (Supplementary Fig. S11 and Supplementary Note 11). In summary, Kar3 is involved in the poleward transport of kinetochores along microtubules. However, unidentified regulators

probably act redundantly with Kar3, because kinetochore transport continued in the majority of cells with *kar3* deletions.

Capture of authentic centromeres in normal cell cycles

To address whether kinetochores on authentic centromeres associate with the microtubule lattice (without regulation by an adjacent *GAL1-10* promoter), we first studied the interaction of kinetochores (marked with Mtw1 and Ctf19, each fused with four tandem copies of GFP) with microtubules after cells were released from nocodazole arrest (Supplementary Fig. S12 and Supplementary Video S11).

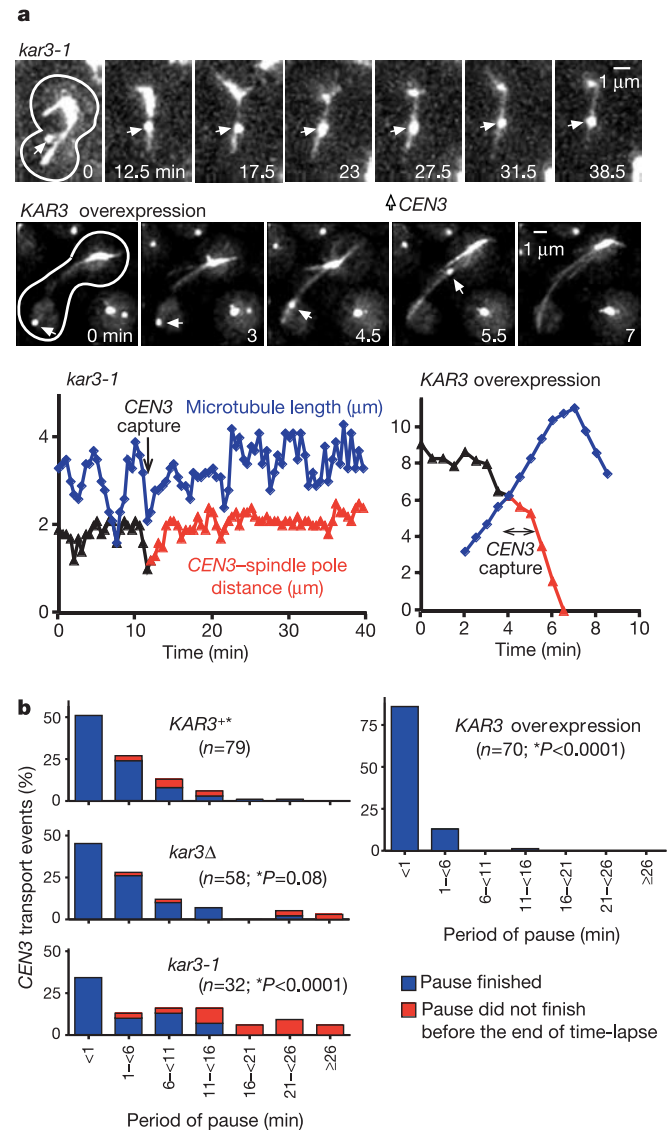


Figure 5 Kar3 kinesin is involved in the poleward transport of kinetochores along the side of microtubules. *KAR3*⁺ cells (T3531, see Fig. 1 legend), cells with the same genotype but including *kar3Δ* (T2864) or *kar3-1* (T3319), and *P_{GAL}-KAR3* cells (T3616) were treated as in Fig. 4a, except that time-lapse images were collected every 30 s for 40 min. **a**, *CEN3* pausing in a *kar3-1* cell (upper panel) and accelerated movement of *CEN3* in a cell overexpressing *KAR3* (lower panel). Scale bar, 1 μm. Graphs indicate microtubule length and *CEN3*-spindle pole distance in the two cells; colours as in Fig. 3a. See also Supplementary Fig. S10 and Supplementary Videos S8–10. **b**, Period of *CEN3* pause. Graphs show the percentage of *CEN3* transport events with the indicated pausing periods (the longest pause in each cell). Pauses that did not finish before the end of the time lapse (red) will have underestimated pause lengths. *n* refers to the number of observed *CEN3* transport events.

Soon after the microtubules had been reformed, several kinetochores moved along the microtubule lattice towards spindle poles.

Second, we released the above cells from α -factor arrest and tried to find individual kinetochore–microtubule interactions during S phase (determined by the emergence of small buds). Kinetochores on a single or a few chromosomes were sometimes found away from other kinetochores that clustered at a spindle pole (Fig. 6a and Supplementary Video S12). These isolated kinetochores seemed to be captured by the side of microtubules (Fig. 6a, 170 s) and transported poleward along microtubules (170–200 s). Such kinetochore association with the side of microtubules was not found in G1 or after establishment of the bipolar spindle.

Third, we used transmission electron microscopy (TEM) to observe nuclear microtubules and to locate centromeres. *CEN3*, 4 and 5 were marked with adjacent *tet* operators bound by TetR–GFP in the same strain. The cells were frozen with high pressure during S phase, and processed for TEM and immuno-gold staining against GFP³⁹. Clusters of two (or more) gold particles were found in the vicinity of microtubules only in cells with GFP-marked *CENs* (that is, with both TetR–GFP and *tet* operators), but not in control cells (with TetR–GFP but without *tet* operators), suggesting that such clusters locate one of the GFP-marked centromeres. In some cells, the clusters were found in the vicinity of SPBs and associated with the side of microtubules (Fig. 6b). For instance, in the right panel of Fig. 6b, one of the *CENs* seems to associate with a microtubule that extends from one of the duplicated but not-yet-separated SPBs (SPB2). Collectively, these results suggest that authentic centromeres associate laterally with microtubules during S phase of an unperturbed cell cycle. (See Supplementary Note 12.)

Discussion

We propose that the following steps occur during kinetochore capture by spindle microtubules in budding yeast (Fig. 6c and Supplementary Fig. S13). During the G1 phase of the cell cycle, kinetochores stay attached to microtubules and are localized in the vicinity of SPBs. Upon centromere DNA replication in early S phase, kinetochores are disassembled and centromeres detach from microtubules. Throughout much of S phase, assembly of a new SPB proceeds in the vicinity of an old SPB that is inherited from the previous cell cycle^{9,40}. When kinetochores are reassembled on both replicated sister centromeres, more nuclear microtubules extend from the old SPB than from the new SPB that is not yet fully operational^{10,11}. Microtubule extension is facilitated by +TIPs (Stu2, Bim1 and Bik1) and a high concentration of RanGTP in the nucleus, and it occurs in various directions (not only towards kinetochores). These microtubules (single microtubules in many cases) capture the kinetochores laterally. Capture requires the CBF3, Ndc80, Mtw1 and Ctf19 kinetochore complexes, but not the Ipl1 or Dam1 complex.

Once captured, kinetochores are transported along microtubules, predominantly towards spindle poles. The kinetochore-loaded Kar3 kinesin motor and other regulators drive this transport. Stu2 localizes at microtubule plus ends and at kinetochores that are not yet captured by microtubules. The amount of Stu2 protein at the microtubule plus end gradually decreases as microtubules shrink. However, after kinetochores have been captured, Stu2 protein is intermittently transported from kinetochores along microtubules towards their plus ends, where newly arrived Stu2 facilitates microtubule rescue (for example, Fig. 4b). In some cases, the plus ends of shrinking microtubules arrive at the kinetochores (for example, the first rescue in Fig. 3a, bottom) before Stu2 has been transported from kinetochores to microtubule plus ends. In such instances, microtubules are still rescued by Stu2, but this rescue occurs at the position of the kinetochore (the amount of Stu2 at the microtubule plus end increases upon rescue; data not shown). Without this rescue, kinetochores would frequently slide off microtubules, because microtubules shrink with a higher velocity than the

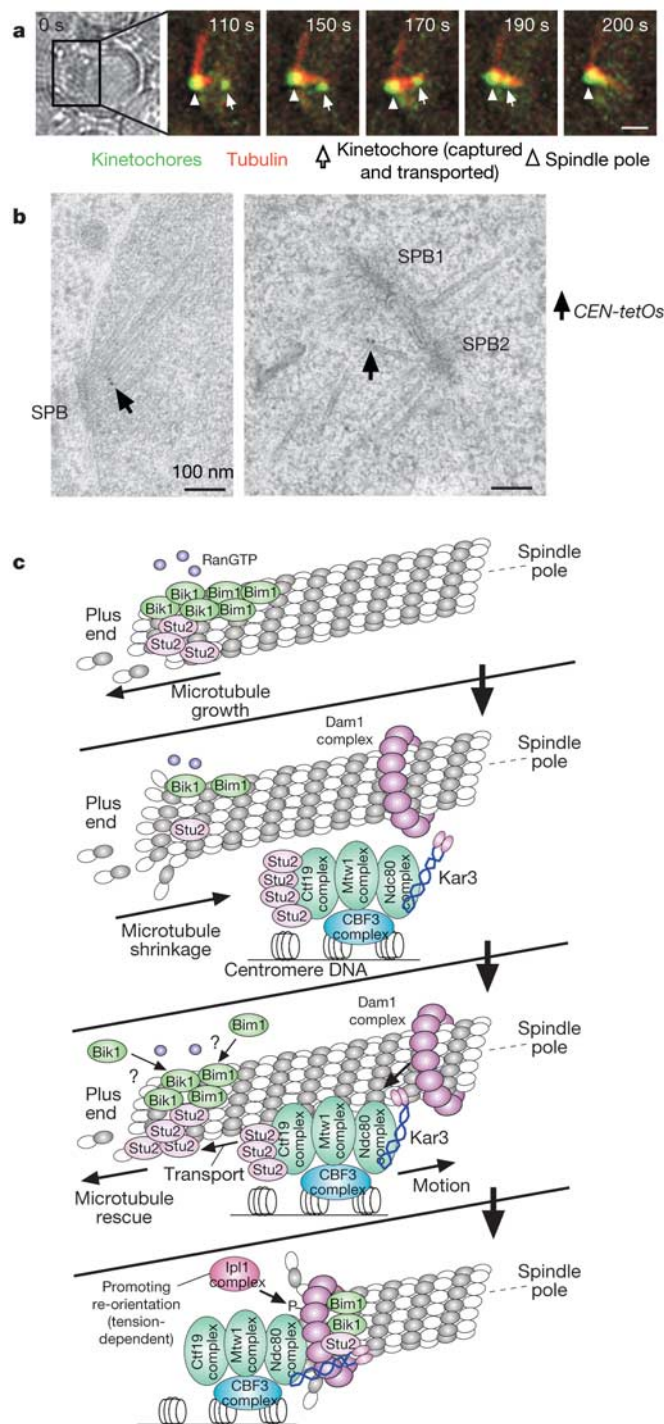


Figure 6 Association of authentic centromeres with microtubules in unperturbed cell cycles. **a**, *MTW1-4GFP CTF19-4GFP YFP-TUB1* cells (T3110) were treated with α -factor for 2.5 h and released to fresh medium. From 60 min after release, GFP and YFP images were acquired every 10 s. Scale bar, 1 μ m. See also Supplementary Video S12. **b**, Transmission electron microscopy. *CEN3-tetOs CEN4-tetOs CEN5-tetOs TetR-GFP* cells (T2841) were fixed by high-pressure freezing when 50% showed bud emergence after release from α -factor arrest. Immuno-gold staining was against GFP. Arrows indicate clusters of two gold particles, marking the position of one of *CEN-tetOs*. Scale bar, 100 nm. **c**, Model for how microtubules interact with kinetochores, and which regulators are involved in kinetochore capture and transport along microtubules. Supplementary Fig. S13 shows a larger scale model.

average poleward motion of captured kinetochores. When kinetochores move along microtubules and approach a spindle pole, both sister kinetochores might attach to microtubules from the same pole (syntelic attachment)^{10,11}.

Later in S phase, the new SPB becomes functional and starts extending nuclear microtubules. Kinetochores change their associated microtubules from the old SPB to those of the new SPB, and continue to switch between the old and new SPBs. The Ipl1 kinase complex promotes this process through phosphorylation of kinetochore components (Dam1, Spc34 and possibly others)^{10,41}. At the end of S phase, SPBs separate to form a bipolar spindle. As sister kinetochores bi-orient and each sister is pulled towards opposite spindle poles, tension is applied on kinetochore–spindle pole connections. When tension is applied, Ipl1-dependent re-orientation of kinetochore–spindle pole connections ceases, leading to stabilization of bi-orientation¹⁸. (See Supplementary Note 13.)

The Dam1 complex is required neither for kinetochore capture nor for transport along the microtubule lattice, but it is crucial for chromosome bi-orientation^{26–29}. We propose that kinetochores use different contact molecules when they associate with microtubules laterally and when sister kinetochores bi-orient on the spindle (Fig. 6c). After kinetochore disassembly upon centromere DNA replication, the Ctf19, Mtw1 and Ndc80 complexes are probably assembled on replicated centromeres (dependent on the CBF3 complex^{17,42–44}) before kinetochores first encounter microtubules, and one or more of the three complexes provides a direct interface with microtubules during lateral attachment to them. In contrast, the Dam1 complex on spindle microtubules is loaded on kinetochores (dependent on Ndc80) only after kinetochores are captured by microtubules^{28,29} (see Supplementary Note 14). The Dam1 complex then acts to stabilize kinetochore–microtubule (perhaps end-on) attachments upon sister kinetochore bi-orientation²⁸. The Dam1 complexes might facilitate kinetochore attachment at the plus end of a dynamic microtubule by forming a ring that encircles the microtubule^{45,46}.

In both vertebrate cells^{2–4} and budding yeast, kinetochores are initially captured by microtubules on their lateral surface. Kinetochore attachment with the side of microtubules was also reported in diatoms⁴⁷. This mechanism might have arisen early in the evolution of eukaryotic cells, because the microtubule lattice can secure initial kinetochore capture by providing a much larger contact surface for this capture compared with microtubule tips. However, to maintain association with kinetochores, the end-on attachment of microtubules might be more stable than the lateral attachment (see Supplementary Note 15). In fact, kinetochores sometimes detach from the microtubule lattice during their poleward transport.

Although kinetochores are captured by microtubules laterally in both budding yeast and vertebrate cells, we notice some important differences. First, microtubules extend from centrosomes preferentially in the direction of chromosomes in *Xenopus* egg extracts⁴⁸. However, in budding yeast, microtubules do not preferentially extend in the direction of kinetochores. Because of the smaller size of its nucleus, yeast might not need such mechanisms. Second, kinetochores slide along the microtubule lattice much faster ($11\text{--}14\ \mu\text{m min}^{-1}$ in newt lung cells⁴⁹) in vertebrate cells than in budding yeast ($0.5\text{--}2.0\ \mu\text{m min}^{-1}$). Whereas dynein is not involved in kinetochore transport along microtubules in *S. cerevisiae*, dynein might regulate fast kinetochore transport in vertebrate cells^{2,50}. If this is the case, eukaryotic cells must have acquired the ability to use cytoplasmic dynein for kinetochore transport after they developed open mitosis. Alternatively, a member of the kinesin-14 family, to which Kar3 belongs, might be involved in fast kinetochore transport in vertebrate cells. (See Supplementary Note 16.)

Our study has revealed mechanisms for kinetochore capture by spindle microtubules in budding yeast. Kinetochore capture is the first crucial step for proper chromosome segregation in all eukaryotic cells. It is likely that, for such fundamental cellular events,

some underlying mechanisms are conserved from yeast to vertebrates, even if modifications were added later in the evolutionary process. □

Methods

Yeast genetics and molecular biology

Yeast strain background (W303), methods for yeast culture, and the TetR–GFP/*tet* operator system were described previously^{10,13,18}. Cells were cultured at 25°C in YP medium containing glucose, unless otherwise stated. See Supplementary Note 17 for strain constructions, more culture conditions and mutant alleles used in this study.

Microscopy

The general procedures for time-lapse fluorescence microscopy were described previously⁷. Time-lapse images were collected every 10 s for 30 min at 23°C (ambient temperature) unless otherwise stated. Using the DeltaVision microscope (Applied Precision), we acquired 3–7 (0.7- μm apart) z-sections, which were subsequently deconvoluted and projected to two-dimensional images using SoftWoRx software (Applied Precision). Samples for TEM and immunostaining were prepared as previously described³⁹. See Supplementary Note 18 for more details.

Analysing dynamics of nuclear microtubules

To analyse microtubule dynamics, we used time-lapse image sequences that were acquired every 10 s for 30 min, unless otherwise stated. To evaluate the length of microtubules and position of centromeres, we took account of the distance along the z-axis as well as distance on a projected image. Error bars in graphs show the 95% confidence interval (Fig. 2a, b) or the standard deviation (Figs 2c and 3b). See Supplementary Note 19 for more details.

Received 18 October 2004; accepted 18 February 2005; doi:10.1038/nature03483.

- McIntosh, J. R., Grishchuk, E. L. & West, R. R. Chromosome–microtubule interactions during mitosis. *Annu. Rev. Cell Dev. Biol.* **18**, 193–219 (2002).
- Rieder, C. L. & Alexander, S. P. Kinetochores are transported poleward along a single astral microtubule during chromosome attachment to the spindle in newt lung cells. *J. Cell Biol.* **110**, 81–95 (1990).
- Hayden, J. H., Bowser, S. S. & Rieder, C. L. Kinetochores capture astral microtubules during chromosome attachment to the mitotic spindle: direct visualization in live newt lung cells. *J. Cell Biol.* **111**, 1039–1045 (1990).
- Merdes, A. & De Mey, J. The mechanism of kinetochore–spindle attachment and polewards movement analyzed in PtK2 cells at the prophase–prometaphase transition. *Eur. J. Cell Biol.* **53**, 313–325 (1990).
- Goshima, G. & Yanagida, M. Establishing biorientation occurs with precocious separation of the sister kinetochores, but not the arms, in the early spindle of budding yeast. *Cell* **100**, 619–633 (2000).
- He, X., Athana, S. & Sorger, P. K. Transient sister chromatid separation and elastic deformation of chromosomes during mitosis in budding yeast. *Cell* **101**, 763–775 (2000).
- Tanaka, T., Fuchs, J., Loidl, J. & Nasmyth, K. Cohesin ensures bipolar attachment of microtubules to sister centromeres and resists their precocious separation. *Nature Cell Biol.* **2**, 492–499 (2000).
- Pearson, C. G., Maddox, P. S., Salmon, E. D. & Bloom, K. Budding yeast chromosome structure and dynamics during mitosis. *J. Cell Biol.* **152**, 1255–1266 (2001).
- Winey, M. & O'Toole, E. T. The spindle cycle in budding yeast. *Nature Cell Biol.* **3**, E23–E27 (2001).
- Tanaka, T. U. *et al.* Evidence that the Ipl1–Sli15 (Aurora kinase–INCENP) complex promotes chromosome bi-orientation by altering kinetochore–spindle pole connections. *Cell* **108**, 317–329 (2002).
- Tanaka, T. U. Chromosome bi-orientation on the mitotic spindle. *Phil. Trans. R. Soc. Lond. B* (in the press); doi:10.1098/rstb.2004.1612.
- Hill, A. & Bloom, K. Genetic manipulation of centromere function. *Mol. Cell Biol.* **7**, 2397–2405 (1987).
- Michaelis, C., Ciosk, R. & Nasmyth, K. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91**, 35–45 (1997).
- Nasmyth, K., Peters, J. M. & Uhlmann, F. Splitting the chromosome: cutting the ties that bind sister chromatids. *Science* **288**, 1379–1385 (2000).
- O'Toole, E. T., Winey, M. & McIntosh, J. R. High-voltage electron tomography of spindle pole bodies and early mitotic spindles in the yeast *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **10**, 2017–2031 (1999).
- Gupta, M. L. Jr *et al.* β -Tubulin C354 mutations that severely decrease microtubule dynamics do not prevent nuclear migration in yeast. *Mol. Biol. Cell* **13**, 2919–2932 (2002).
- McAins, A. D., Tytell, J. D. & Sorger, P. K. Structure, function, and regulation of budding yeast kinetochores. *Annu. Rev. Cell Dev. Biol.* **19**, 519–539 (2003).
- Dewar, H., Tanaka, K., Nasmyth, K. & Tanaka, T. U. Tension between two kinetochores suffices for their bi-orientation on the mitotic spindle. *Nature* **428**, 93–97 (2004).
- Akhmanova, A. & Hoogenraad, C. C. Microtubule plus-end-tracking proteins: mechanisms and functions. *Curr. Opin. Cell Biol.* **17**, 47–54 (2005).
- Corbett, A. H. & Silver, P. A. Nucleocytoplasmic transport of macromolecules. *Microbiol. Mol. Biol. Rev.* **61**, 193–211 (1997).
- Hetzler, M., Gruss, O. J. & Mattaj, I. W. The Ran GTPase as a marker of chromosome position in spindle formation and nuclear envelope assembly. *Nature Cell Biol.* **4**, E177–E184 (2002).
- Quimby, B. B. & Dasso, M. The small GTPase Ran: interpreting the signs. *Curr. Opin. Cell Biol.* **15**, 338–344 (2003).
- Wilde, A. *et al.* Ran stimulates spindle assembly by altering microtubule dynamics and the balance of motor activities. *Nature Cell Biol.* **3**, 221–227 (2001).
- Carazo-Salas, R. E., Gruss, O. J., Mattaj, I. W. & Karsenti, E. Ran-GTP coordinates regulation of microtubule nucleation and dynamics during mitotic–spindle assembly. *Nature Cell Biol.* **3**, 228–234 (2001).
- Biggins, S. *et al.* The conserved protein kinase Ipl1 regulates microtubule binding to kinetochores in budding yeast. *Genes Dev.* **13**, 532–544 (1999).
- Cheeseman, I. M., Enquist-Newman, M., Muller-Reichert, T., Drubin, D. G. & Barnes, G. Mitotic

- p spindle integrity and kinetochore function linked by the Duo1p/Dam1p complex.
- J. Cell Biol.*
- 152**
- , 197–212 (2001).
27. He, X., Rines, D. R., Espelin, C. W. & Sorger, P. K. Molecular analysis of kinetochore-microtubule attachment in budding yeast. *Cell* **106**, 195–206 (2001).
 28. Janke, C., Ortiz, J., Tanaka, T. U., Lechner, J. & Schiebel, E. Four new subunits of the Dam1–Duo1 complex reveal novel functions in sister kinetochore biorientation. *EMBO J.* **21**, 181–193 (2002).
 29. Li, Y. *et al.* The mitotic spindle is required for loading of the DASH complex onto the kinetochore. *Genes Dev.* **16**, 183–197 (2002).
 30. Maddox, P. S., Bloom, K. S. & Salmon, E. D. The polarity and dynamics of microtubule assembly in the budding yeast *Saccharomyces cerevisiae*. *Nature Cell Biol.* **2**, 36–41 (2000).
 31. Tirnauer, J. S., O'Toole, E., Berrueta, L., Bierer, B. E. & Pellman, D. Yeast Bim1p promotes the G1-specific dynamics of microtubules. *J. Cell Biol.* **145**, 993–1007 (1999).
 32. van Breugel, M., Drechsel, D. & Hyman, A. Stu2p, the budding yeast member of the conserved Dis1/XMAP215 family of microtubule-associated proteins is a plus end-binding microtubule destabilizer. *J. Cell Biol.* **161**, 359–369 (2003).
 33. Lin, H. *et al.* Polyploids require Bik1 for kinetochore–microtubule attachment. *J. Cell Biol.* **155**, 1173–1184 (2001).
 34. Carvalho, P., Gupta, M. L. Jr, Hoyt, M. A. & Pellman, D. Cell cycle control of kinesin-mediated transport of Bik1 (CLIP-170) regulates microtubule stability and dynein activation. *Dev. Cell* **6**, 815–829 (2004).
 35. Hildebrandt, E. R. & Hoyt, M. A. Mitotic motors in *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* **1496**, 99–116 (2000).
 36. Saunders, W. S., Koshland, D., Eshel, D., Gibbons, I. R. & Hoyt, M. A. *Saccharomyces cerevisiae* kinesin- and dynein-related proteins required for anaphase chromosome segregation. *J. Cell Biol.* **128**, 617–624 (1995).
 37. Meluh, P. B. & Rose, M. D. *KAR3*, a kinesin-related gene required for yeast nuclear fusion. *Cell* **60**, 1029–1041 (1990).
 38. Maddox, P. S., Stemple, J. K., Satterwhite, L., Salmon, E. D. & Bloom, K. The minus end-directed motor Kar3 is required for coupling dynamic microtubule plus ends to the cortical shmoo tip in budding yeast. *Curr. Biol.* **13**, 1423–1428 (2003).
 39. Muller-Reichert, T. *et al.* Analysis of the distribution of the kinetochore protein Ndc10p in *Saccharomyces cerevisiae* using 3-D modeling of mitotic spindles. *Chromosoma* **111**, 417–428 (2003).
 40. Adams, I. R. & Kilmartin, J. V. Spindle pole body duplication: a model for centrosome duplication? *Trends Cell Biol.* **10**, 329–335 (2000).
 41. Cheeseman, I. M. *et al.* Phospho-regulation of kinetochore-microtubule attachments by the Aurora kinase Ipl1p. *Cell* **111**, 163–172 (2002).
 42. Westermann, S. *et al.* Architecture of the budding yeast kinetochore reveals a conserved molecular core. *J. Cell Biol.* **163**, 215–222 (2003).
 43. Nekrasov, V. S., Smith, M. A., Peak-Chew, S. & Kilmartin, J. V. Interactions between centromere complexes in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **14**, 4931–4946 (2003).
 44. De Wulf, P., McAnish, A. D. & Sorger, P. K. Hierarchical assembly of the budding yeast kinetochore from multiple subcomplexes. *Genes Dev.* **17**, 2902–2921 (2003).
 45. Miranda, J. L., De Wulf, P., Sorger, P. & Harrison, S. C. The yeast DASH complex forms closed rings on microtubules. *Nature Struct. Mol. Biol.* **12**, 138–143 (2005).
 46. Westermann, S. *et al.* Formation of a dynamic kinetochore-microtubule interface through assembly of the Dam1 ring complex. *Mol. Cell* **17**, 277–290 (2005).
 47. Pickett-Heaps, J. D. Cell division in diatoms. *Int. Rev. Cytol.* **128**, 63–107 (1991).
 48. Carazo-Salas, R. E. & Karsenti, E. Long-range communication between chromatin and microtubules in *Xenopus* egg extracts. *Curr. Biol.* **13**, 1728–1733 (2003).
 49. Alexander, S. P. & Rieder, C. L. Chromosome motion during attachment to the vertebrate spindle: initial saltatory-like behavior of chromosomes and quantitative analysis of force production by nascent kinetochore fibers. *J. Cell Biol.* **113**, 805–815 (1991).
 50. King, J. M., Hays, T. S. & Nicklas, R. B. Dynein is a transient kinetochore component whose binding is regulated by microtubule attachment, not tension. *J. Cell Biol.* **151**, 739–748 (2000).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank J. R. Swedlow, M. J. R. Stark, A. Gartner, M. A. Hoyt, A. Desai, P. R. Clarke, P. D. Andrews and members of the Tanaka laboratory for discussions and for reading the manuscript; T. Hyman, K. Nasmyth, I. W. Mattaj, E. Karsenti, F. Uhlmann and J. Ellenberg for discussions; J.-F. Maure, N. Rachidi and M. J. R. Stark for sharing their unpublished data; Y. Kitamura, S. Swift, M. Romao and G. Keir for technical help; F. Wheatley and the media kitchen for media preparation; M. A. Hoyt, D. Pellman, R. Ciosk, F. Uhlmann, K. Nasmyth, T. C. Huffaker, J. V. Kilmartin, E. Schiebel, S. Biggins, C. S. M. Chan, I. M. Cheeseman, G. Barnes, R. Tsien, S. J. Elledge, J. Lechner, A. H. Corbett, P. A. Silver, P. K. Sorger, X. He, A. F. Straight, M. D. Rose, V. Doye, F. Severin, I. Ouspenski, K. Bloom, T. Nishimoto, J. E. Haber, T. N. Davis, EUROSCARF and the Yeast Resource Center for reagents. This work was supported by The Wellcome Trust, Cancer Research UK and the EMBO Young Investigator Program.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.U.T (t.tanaka@dundee.ac.uk).

No high-mass protostars in the silhouette young stellar object M17-SO1

Shigeyuki Sako^{1,2,3}, Takuya Yamashita^{1,2}, Hirokazu Kataza⁴, Takashi Miyata³, Yoshiko K. Okamoto⁵, Mitsuhiro Honda^{1,2,4}, Takuya Fujiyoshi², Hiroshi Terada², Takeshi Kamazaki³, Zhibo Jiang⁶, Tomoyuki Hanawa⁷ & Takashi Onaka¹

¹Department of Astronomy, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan

²Subaru Telescope, National Astronomical Observatory of Japan, National Institutes of Natural Sciences, 650 North A'ohoku Place, Hilo, Hawaii 96720-2700, USA

³Institute of Astronomy, The University of Tokyo, 2-21-1 Osawa, Mitaka, Tokyo 181-0015, Japan

⁴Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency, 3-1-1 Yoshinodai, Sagami-hara, Kanagawa 229-8510, Japan

⁵Institute of Astrophysics & Planetary Sciences, Ibaraki University, 2-1-1 Bunkyo, Mito, Ibaraki 310-8512, Japan

⁶Purple Mountain Observatory, National Astronomical Observatories, Chinese Academy of Sciences, Nanjing 210008, China

⁷Centre for Frontier Science, Chiba University, Yayoi-cho, Inage-ku, Chiba 263-8522, Japan

The birth of very massive stars is not well understood^{1–3}, in contrast to the formation process of low-mass stars like our Sun^{4,5}. It is not even clear that massive stars can form as single entities; rather, they might form through the mergers of smaller ones born in tight groups^{6,7}. The recent claim of the discovery of a massive protostar in M17 (a nearby giant ionized region) forming through the same mechanism as low-mass stars⁸ has therefore generated considerable interest. Here we show that this protostar has an intermediate mass of only 2.5 to 8 solar masses ($M_{\odot}$), contrary to the earlier claim of $20M_{\odot}$ (ref. 8). The surrounding circumstellar envelope contains only $0.09M_{\odot}$ and a much more extended local molecular cloud has $4\text{--}9M_{\odot}$.

M17-silhouette object 1 (M17-SO1; Fig. 1a) is the most prominent silhouette young stellar object (YSO) discovered in our survey, which is being carried out independently of the work reported in ref. 8. The silhouette against the background light of the M17 H II region consists of an opaque inner region ('inner envelope') and a less-opaque region outside a 2,200 AU radius ('outer envelope'). Both envelopes are flared in shape and concentric around a reddened central source associated with bipolar nebulae, as seen in a multi-colour near-infrared image (Fig. 2a). The appearance is similar to infrared and optical images of low- or intermediate-mass class I YSOs in an early evolutionary stage surrounded by disks and envelopes, which show flaring structures of the scattered light in observations by the Hubble Space Telescope and ground-based telescopes⁹. The present Br γ image and the interferometric observations in centimetre wavelengths¹⁰ contradict evidence of compact H II regions at M17-SO1, and no mid-infrared sources like Orion BN/KL¹¹ and W51¹² are detected in the present 12.8- μ m image shown in Fig. 1b. The absence of a compact H II region and no infrared emission indicate that the central star is not producing ionizing photons and hence its mass is no more than $8M_{\odot}$ (ref. 3). The bipolar scattered light image is also similar to those of several class I objects of a few solar masses^{13,14} with respect to appearance, size and surface brightness^{15,16} (Fig. 2b). An illustration of the features detected in the infrared images is shown in Fig. 2c. Although it was claimed in ref. 8 that the mass of the central star is about $20M_{\odot}$ with a distance of 2.2 kpc for M17, the present new observations of M17-SO1 clearly indicate that it is a low- or intermediate-mass star of less than $8M_{\odot}$.

The silhouette image against the uniform background light source gives an unambiguous column density distribution of the dusty component as shown in Fig. 1. The total mass of the inner and outer envelopes is estimated to be $0.09M_{\odot}$ by integrating the column density map within the envelope area. This mass lies in a typical envelope mass range for low- and intermediate-mass class I objects of $\sim(0.02\text{--}0.3)M_{\odot}$ (refs 4, 9). The mass of the local molecular cloud, in which the silhouette envelope is embedded, is estimated to be $4M_{\odot}$ by integrating the column density map as well as that of the envelopes. The cloud mass is estimated to be $9M_{\odot}$ also from the total flux of ^{13}CO (Fig. 3a) using the physical relationship between the flux and the column density as described in the Fig. 3a legend. The agreement within a factor of a few between the two mass estimates supports its relevance, despite the uncertainties due to possible CO depletion¹⁷ or in the employed conversion factor. On the other hand, ref. 8 identified it as a massive disk of more than $100M_{\odot}$ on the basis of their ^{13}CO observations using an empirical relationship between the ^{12}CO flux and the column density with a typical abundance ratio of $^{12}\text{CO}/^{13}\text{CO}$. The velocity structure of the local molecular cloud appears to rotate around M17-SO1 (Fig. 3b). The position-velocity diagram along the disk mid-plane is shown in Supplementary Information. The velocity difference between the positions separated by $\sim 15,000$ AU is ~ 1.0 km s⁻¹, indicating that the total system mass is $\sim 8M_{\odot}$ if the rotation is keplerian. This mass also agrees well with the present estimates, but is incompatible with a massive star and disk⁸.

The differences in mass estimate between the present study and ref. 8 are summarized in Table 1. A major discrepancy is attributed to the relationships used. The prime reason that we claim the relevance of the present results is the consistency between the mass estimates from the molecular emission and the dust extinction without assuming the dust temperature. Part of the difference is also due to the assumed distance from the Sun to M17. The present study adopts 1.5 kpc, which has been frequently used in recent studies of M17^{18–20}, whereas a larger value of 2.2 kpc was used in ref. 8. The difference in the assumed distance accounts for a factor of 2 of the difference in the mass estimates. We conclude that M17-SO1 is a low- or intermediate-mass class I object with a flared envelope of $0.09M_{\odot}$, which is surrounded by a molecular cloud of $(4\text{--}9)M_{\odot}$, and that there is no evidence for ongoing disk accretion onto massive stars, as claimed by ref. 8.

The present silhouette images reveal the column density distribution of the circumstellar dust from the region near the central star to outer radii of more than 10,000 AU, which are not obtained from the emission features. It is difficult to derive the density distribution and study the star/disk formation process solely from the scattered

Table 1 Differences in mass estimates between the present study and ref. 8

Method of estimate	The present study		Ref. 8 2.2 kpc*
	1.5 kpc*	2.2 kpc*	
Mass of the central star ($M_{\odot}$)			
Ionized region	<8 (no ionized region)		
Velocity structure of ^{13}CO	8†	12†	15
Flared structure	>2.5	>4.5	
Photometry of the central source	Regarded as scattered light		20
Total mass of the envelopes and the local molecular cloud ($M_{\odot}$)			
Total flux of ^{13}CO	9‡	19‡	>100§
2 μ m dust extinction	4	9	
Velocity structure of ^{13}CO	8†	12†	

Mass estimates using a distance to M17 of 2.2 kpc are also listed, together with the 1.5 kpc case in the column for the present study. The present mass estimates differ greatly from that of ref. 8, even for the 2.2 kpc case. A major discrepancy can be attributed to the employed relationships themselves.

*Assumed distance.

†Total mass of the central star, local molecular cloud, envelopes and circumstellar disk.

‡Derived using physical relationship. Details are described in Fig. 3a legend.

§Derived using empirical relationship. Details are described in the text.

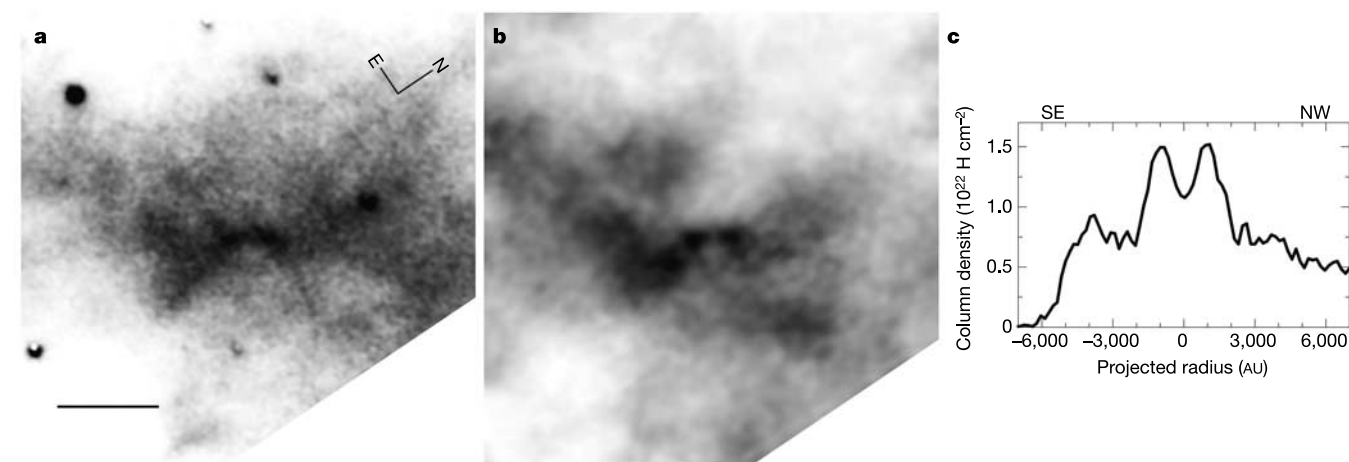


Figure 1 Silhouette features of M17-SO1. **a**, Silhouette image of M17-SO1 against the Br γ (2.166 μm) emission. This image was obtained by the IRCS²⁵ with the AO²⁶ system at the 8.2-m Subaru telescope on 23 May 2003. It is constructed by subtracting the normalized K_{cont} (2.269 μm) narrow band filter (NBF) image from the Br γ NBF image. As the continuum emission has been subtracted, the image shows a distribution of dusty component, free from the emission from the central compact source and the scattered light. The displayed area is 14.8'' $\times$ 14.8'' and the length of the bar indicates 3.3'' or 5,000 AU. The north and east are shown by bars. The position angle of the mid-plane is 146°. The spatial resolution is about 0.2''. The flared envelope is clearly seen at the centre of the image. The masses of the inner and outer envelopes are estimated to be 0.006 $M_{\odot}$ and 0.08 $M_{\odot}$, respectively, by integrating the column density map shown in Fig. 3a within the envelope area and assuming a gas to dust mass ratio of 100:1. **b**, 12.8- μm

mid-infrared image of M17-SO1. This image was taken with the Cooled Mid-Infrared Camera and Spectrometer (COMICS²⁷) at the Subaru telescope on 15 June 2003. The silhouette structures are consistent, in their shape and size, with those in the Br γ silhouette image shown in **a**. The central compact source seen in the H and K' band images is not found in this image. **c**, Column density profile of the outer and inner envelopes, $N_{z=0}(x)$, derived from the silhouette image in **a**. The horizontal axis indicates the offset projected radius from the position of the central source along the disk mid-plane. Since the envelopes are surrounded by the local molecular cloud with a column density of $\sim 3.5 \times 10^{22}$ H atoms cm^{-2} , the peak of the column density in this region is $\sim 5 \times 10^{22}$ H atoms cm^{-2} , which approximately corresponds to the estimate in ref. 8, $\sim 6 \times 10^{22}$ H atoms cm^{-2} . $N_{z=0}(x)$ shows artificial structure at radii less than 500 AU because the scale height is unresolved at these radii.

light image because the image depends on the position of the light source. The flared shape of the envelope would be gravitationally related to the mass of the central star. The vertical scale height (e-folding length) of the column density distribution is 1,900 AU at the outer boundary of the envelope at 4,000 AU radius, where it approximately reflects a scale height of the volume density

distribution of the envelope if the envelope has an axisymmetric structure and a truncated outer boundary. The scale height (H) of the volume density distribution follows the formula:

$$H(r) = [(kT)/(\mu m_{\text{H}} G M_{\star})]^{1/2} r^{1.5} \quad (1)$$

where k is the Boltzmann constant, T is the gas kinetic temperature,

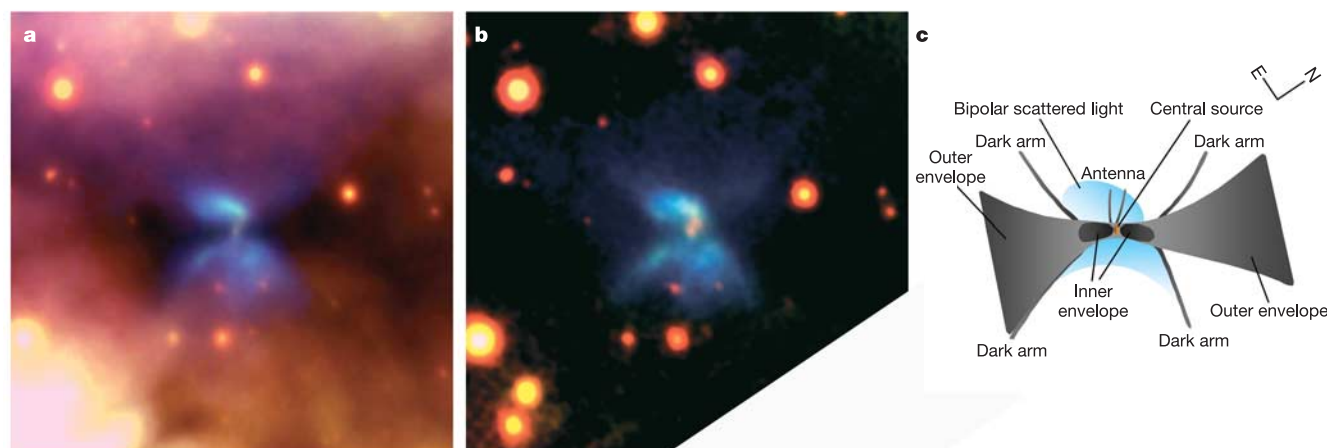


Figure 2 Emission features and illustration of M17-SO1. **a**, J (1.3 μm), H (1.6 μm), K' (2.1 μm) pseudo-colour image. The image was taken in follow-up observations with the IRCS and AO system on 15 August 2003. The AO corrected spatial resolution is 0.36'', 0.27'' and 0.14'' in J, H and K' bands, respectively. The stellar-like central source is located at 18 h 20 min 26.18 s, $-16^{\circ} 12' 10.2''$ (J2000). The orientation and the displayed size are the same as in Fig. 1a. The envelope size is consistent with those of Taurus class I envelopes ranging from 1,500 AU to 8,000 AU defined by the full-width at half-maximum of the molecular line emissions^{28, 29}. An elongated compact infrared source is found near the symmetric centre of the envelope in the H and K' band images. This compact infrared source is not a direct stellar image itself, but rather scattered light

due to the elongation. A faint point source is found in the elongated source in the K' band image. **b**, Reconstructed JHK' composite image of M17-SO1 if it was located in a region without background light. This image is produced by subtracting the background component, which is estimated with the optical depth shown in Fig. 3a. The surface brightness is 150 and 110 $\mu\text{Jy arcsec}^{-2}$ in the J band at the peaks of the scattered light in the northeast and southwest sides, respectively. **c**, Illustration of the features detected in the net Br γ image and the JHK composite image of **a**. Two 'antennae' have no counterparts in other observational features. The structures remind us of a dust shell as well as the bipolar dust walls observed as the four dark arms. They may indicate the presence of a highly collimated jet sweeping up circumstellar dust.

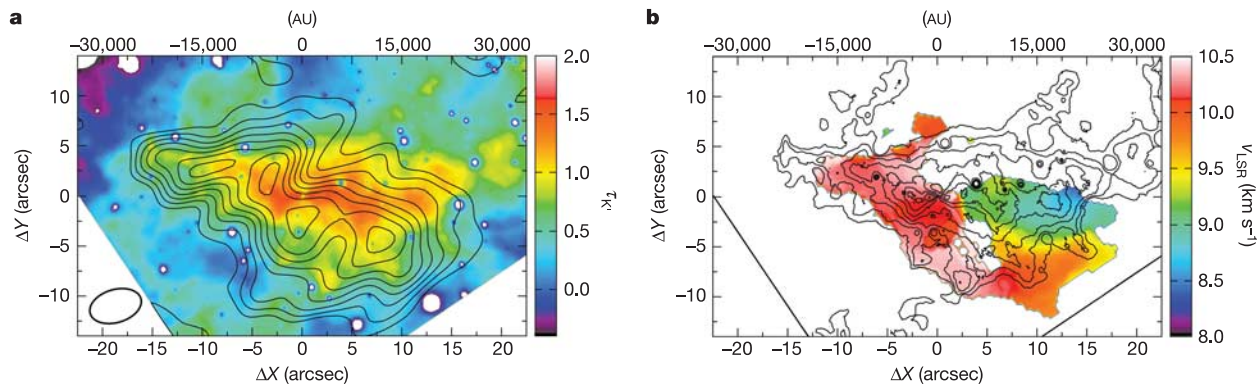


Figure 3 ^{13}CO 1–0 line images of M17-SO1. The images were obtained with the Nobeyama millimetre array on December 2003 and April 2004. The beam size is $3.3'' \times 5.3''$, as shown at the left lower corner in **a**. The flared envelope and star/disk system of M17-SO1 are located in the local molecular cloud of size $45,000 \text{ AU} \times 22,000 \text{ AU}$. **a**, ^{13}CO 1–0 integrated intensity map is drawn with contours starting from $0.33 \text{ Jy km}^{-1} \text{ beam}^{-1}$ in steps of $0.33 \text{ Jy km}^{-1} \text{ beam}^{-1}$. Pseudo-colour map shows the optical depth (τ) in the K' band, which is not correct near the stellar emission including the scattered light of M17-SO1. The mass of the local molecular cloud including the flared envelope is estimated to be $9M_{\odot}$ from the total intensity of the ^{13}CO 1–0 line emission of 39 Jy km^{-1} , when the excitation temperature and abundance for the ^{13}CO gas are 30 K and 10^{-6} , respectively, and the optical depth of the ^{13}CO 1–0 line

μ is the mean molecular weight, m_{H} is the mass of the hydrogen atom, G is the gravitational constant, $M_{\star}$ is the mass of the central star, and r is the disk equatorial radius²¹, when the thermal pressure gradient in the envelope is in dynamical equilibrium with the vertical component of the stellar gravity. The scale height at the outer boundary yields a relation $M_{\star}/T \approx 0.08M_{\odot} \text{ K}^{-1}$. If we take the observed CO excitation temperature in this region of $30\text{--}35 \text{ K}$ (refs 22, 23) as the gas kinetic temperature, $M_{\star}$ becomes $2.5M_{\odot}$. In practice, the turbulent motion and/or the magnetic field should provide internal pressure in addition to the thermal pressure that acts to resist gravity. When the envelope is vertically infalling, the scale height appears larger than that predicted with equation (1). Therefore, $2.5M_{\odot}$ is a lower limit of the stellar mass because of the neglect of the non-thermal effects and infall motion. The flared shape of the silhouette suggests that it is an intermediate-mass star with more than $2.5M_{\odot}$.

Figure 1c shows the column density profile along the projected mid-plane, $N_{z=0}(x)$, derived from the map of optical depth. This does not reflect a true column density distribution within 500 AU radius because the disk thickness is not resolved there. The highly extinguished stellar light and the lack of silhouette at the centre suggest that a geometrically thin but optically thick disk with a radius of a few hundred AU is located around the central star. An interesting feature is that $N_{z=0}(x)$ becomes flat or increasing outwards beyond $2,000 \text{ AU}$. This profile would not happen if the volume density were monotonically decreasing outwards in the outer envelope, because the optical length along the line of sight would decrease. This may suggest the existence of a gap, that is, a local minimum, of the volume density distribution between the inner and outer envelopes.

Four long dark 'arms' seen in Fig. 1a are new features found in the present observations. They tightly enclose the blue hourglass nebula shown in Fig. 2b and are clearly separated from the flared envelope structure. This kind of hourglass or bipolar structure is typical of a scattering nebula associated with a molecular outflow/YSO system seen edge-on. The geometrical appearance indicates that the four arms are the shells swept up by a bipolar outflow, whose presence is confirmed by the present observations. The most important feature is that the four arms are detached from the envelope, which

is 0.5 , which is estimated with the optical depth in the K' band and the ^{13}CO velocity dispersion of 1 km s^{-1} . The large velocity dispersion implies that turbulent motion is dominant in the local molecular cloud, but it is not clear in the outer and inner envelopes. **b**, Intensity weighted mean velocity map. Velocities (with respect to the local standard of rest) between $v_{\text{LSR}} = 8 \text{ km s}^{-1}$ and 10.5 km s^{-1} are indicated in pseudo-colour. The contours show the optical depth in the K' band, which start from 0.6 in steps of 0.15 . The velocity structure appears to rotate around M17-SO1, but the exact relationship of the ^{13}CO to the M17-SO1 system is not entirely clear because the velocity field of the ^{13}CO is complicated and the centre of the rotating motion is shifted from the centre of M17-SO1.

indicates that the outflow is not directly interacting with the envelope at all. It is not compatible with a picture in which the molecular outflow is an important agent for evacuating the envelope and/or forming its flattened shape²⁴. The outer envelope has only $0.08M_{\odot}$ and is probably smaller than the mass lost by the outflow, typically a few tenths of a solar mass⁴. The majority of the infalling envelope mass is taken into the star/disk system through the flattened envelope without interaction with the outflow in the class I phase. □

Received 20 September 2004; accepted 11 February 2005; doi:10.1038/nature03471.

- Yorke, H. W. & Sonnhalter, C. On the formation of massive stars. *Astrophys. J.* **569**, 846–862 (2002).
- Wolfire, M. G. & Cassinelli, J. P. Conditions for the formation of massive stars. *Astrophys. J.* **319**, 850–867 (1987).
- Stahler, S. W., Palla, F. & Ho, P. T. P. in *Protostars and Planets IV* (eds Mannings, V., Boss, A. P. & Russell, S.) 327–351 (Univ. Arizona Press, Tucson, 2000).
- Mundy, L. G., Looney, L. W. & Welch, W. J. in *Protostars and Planets IV* (eds Mannings, V., Boss, A. P. & Russell, S.) 355–376 (Univ. Arizona Press, Tucson, 2000).
- Calvet, N., Hartmann, L. & Strom, S. E. in *Protostars and Planets IV* (eds Mannings, V., Boss, A. P. & Russell, S.) 377–399 (Univ. Arizona Press, Tucson, 2000).
- Bonnell, I. A., Bate, M. R. & Zinnecker, H. On the formation of massive stars. *Mon. Not. R. Astron. Soc.* **298**, 93–102 (1998).
- Bonnell, I. A., Vine, S. G. & Bate, M. R. Massive star formation: nurture, not nature. *Mon. Not. R. Astron. Soc.* **349**, 735–741 (2004).
- Chini, R. *et al.* The formation of a massive protostar through the disk accretion of gas. *Nature* **429**, 155–157 (2004).
- Padgett, D. L. *et al.* Hubble space telescope/NICMOS imaging of disks and envelopes around very young stars. *Astron. J.* **117**, 1490–1504 (1999).
- Felli, M., Massi, M. & Churchwell, E. A high-resolution study of M17 at 1.3, 2, 6, and 21 CM. *Astron. Astrophys.* **136**, 53–64 (1984).
- Shuping, R. Y., Morris, M. & Bally, J. A new mid-infrared map of the BN/KL region using the Keck telescope. *Astron. J.* **128**, 363–374 (2004).
- Okamoto, Y. K. *et al.* High-resolution mid-infrared imaging and spectroscopic observations of a massive star forming region W51 IRS 2. *Astrophys. J.* **553**, 254–266 (2001).
- Ohashi, N., Hayashi, M., Ho, P. T. P. & Momose, M. Interferometric imaging of IRAS 04368+2557 in the L1527 molecular cloud core: A dynamically infalling envelope with rotation. *Astrophys. J.* **475**, 211–223 (1997).
- Ohashi, N. *et al.* Rotation in the protostellar envelopes around IRAS 04169+2702 and IRAS 04365+2535: The size scale for dynamical collapse. *Astrophys. J.* **488**, 317–329 (1997).
- Park, S. & Kenyon, S. J. Extended near-infrared emission from candidate protostars in the Taurus-Auriga molecular cloud. *Astron. J.* **123**, 3370–3379 (2002).
- Tamura, M., Gatley, L., Waller, W. & Werner, M. W. Two micron morphology of candidate protostars. *Astrophys. J.* **374**, L25–L28 (1991).
- Thi, W. F. *et al.* H₂ and CO emission from disks around T Tauri and Herbig Ae pre-main-sequence stars and from debris disks around young stars: Warm and cold circumstellar gas. *Astrophys. J.* **561**, 1074–1094 (2001).
- Jiang, Z. *et al.* Deep near-infrared survey toward the M17 region. *Astrophys. J.* **577**, 245–259 (2002).

19. Hanson, M. M., Howarth, I. D. & Conti, P. S. The young massive stellar objects of M17. *Astrophys. J.* **489**, 698–718 (1997).
20. Nielbock, M., Chini, R., Jütte, M. & Manthey, E. High mass Class I sources in M 17. *Astron. Astrophys.* **377**, 273–284 (2001).
21. Hartmann, L. *Accretion Processes in Star Formation* (Cambridge Univ. Press, New York, 1998).
22. Wilson, C. D., Howe, J. E. & Balogh, M. L. The large-scale $J = 3 \rightarrow 2$ and $J = 2 \rightarrow 1$ CO emission from M17 and its implications for extragalactic CO observations. *Astrophys. J.* **517**, 174–187 (1999).
23. Howe, J. E. *et al.* Extended [C I] and ^{13}CO ($5 \rightarrow 4$) emission in M17SW. *Astrophys. J.* **539**, 137–141 (2000).
24. Nakano, T., Hasegawa, T. & Norman, C. The mass of a star formed in a cloud core: Theory and its application to the Orion A cloud. *Astrophys. J.* **450**, 183–195 (1995).
25. Kobayashi, N. *et al.* IRCs: infrared camera and spectrograph for the Subaru Telescope. *Proc. SPIE* **4008**, 1056–1066 (2000).
26. Takami, H. *et al.* Performance of Subaru adaptive optics system and the scientific results. *Proc. SPIE* **4839**, 21–31 (2003).
27. Kataza, H. *et al.* COMICS: the cooled mid-infrared camera and spectrometer for the Subaru telescope. *Proc. SPIE* **4008**, 1144–1152 (2000).
28. Hogerheide, M. R., van Dishoeck, E. F., Blake, G. A. & van Langevelde, H. J. Envelope structure on 700 AU scales and the molecular outflows of low-mass young stellar objects. *Astrophys. J.* **502**, 315–336 (1998).
29. Ohashi, N., Hayashi, M., Kawabe, R. & Ishiguro, M. The Nobeyama Millimeter Array survey of young stellar objects associated with the Taurus molecular cloud. *Astrophys. J.* **466**, 317–337 (1996).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements This Letter is based on data collected at the Subaru Telescope and the Nobeyama Radio Observatory, which are operated by the National Astronomical Observatory of Japan.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.S. (sako@ioa.s.u-tokyo.ac.jp).

Chronology of the early Solar System from chondrule-bearing calcium-aluminium-rich inclusions

Alexander N. Krot¹, Hisayoshi Yurimoto², Ian D. Hutcheon³ & Glenn J. MacPherson⁴

¹Hawai'i Institute of Geophysics & Planetology, School of Ocean & Earth Science & Technology, University of Hawai'i at Manoa, Honolulu, Hawaii 96822, USA

²Department of Earth and Planetary Sciences, Tokyo Institute of Technology, Meguro, Tokyo 152-8551, Japan

³Institute of Geophysics & Planetary Physics and G.T. Seaborg Institute, Lawrence Livermore National Laboratory, Livermore, California 94451, USA

⁴Smithsonian Institution, Department of Mineral Sciences, NHB 119, Washington DC 20560, USA

Chondrules and Ca-Al-rich inclusions (CAIs) are high-temperature components of meteorites that formed during transient heating events in the early Solar System. A major unresolved issue is the relative timing of CAI and chondrule formation^{1–4}. From the presence of chondrule fragments in an igneous CAI, it was concluded that some chondrules formed before CAIs (ref. 5). This conclusion is contrary to the presence of relict CAIs inside chondrules^{6–10}, as well as to the higher abundance of ^{26}Al in CAIs¹¹; both observations indicate that CAIs pre-date chondrules by 1–3 million years (Myr). Here we report that relict chondrule material in the Allende meteorite, composed of olivine and low-calcium pyroxene, occurs in the outer portions of two CAIs and is ^{16}O -poor ($\Delta^{17}\text{O} \approx -1\text{‰}$ to -5‰). Spinel and diopside in the CAI cores are ^{16}O -rich ($\Delta^{17}\text{O}$ up to -20‰), whereas diopside in their outer zones, as well as melilite and anorthite, are ^{16}O -depleted ($\Delta^{17}\text{O} = -8\text{‰}$ to 2‰). Both chondrule-bearing CAIs are ^{26}Al -poor with initial $^{26}\text{Al}/^{27}\text{Al}$ ratios of

$(4.7 \pm 1.4) \times 10^{-6}$ and $<1.2 \times 10^{-6}$. We conclude that these CAIs had chondrule material added to them during a re-melting episode ~ 2 Myr after formation of CAIs with the canonical $^{26}\text{Al}/^{27}\text{Al}$ ratio of 5×10^{-5} .

Mineralogical, chemical and isotopic data suggest that refractory inclusions formed in an ^{16}O -rich gaseous reservoir ($\Delta^{17}\text{O}_{\text{SMOW}} \approx -20\text{‰}$) at high ambient temperatures (near or above the condensation temperatures of forsterite; $\sim 1,375$ K at a total pressure of 10^{-4} bar), and were subsequently isolated (physically or kinetically) from reactions with the high temperature solar nebula gas¹. (Here $\Delta^{17}\text{O} = \delta^{17}\text{O} - 0.52 \times \delta^{18}\text{O}$; $\delta^{17,18}\text{O} = [(^{17,18}\text{O}/^{16}\text{O})_{\text{sample}} / (^{17,18}\text{O}/^{16}\text{O})_{\text{SMOW}} - 1] \times 1,000$, where SMOW is Standard Mean Ocean Water.) Evaporation and condensation are believed to have been the dominant processes during formation of refractory inclusions¹. Subsequently, some CAIs (called 'igneous') experienced extensive melting accompanied by evaporation¹². Both igneous and non-igneous CAIs are surrounded by ^{16}O -rich multilayered rims (called 'Wark-Lovering' rims), with the outermost layers, composed of Al-diopside and forsterite, probably formed by condensation¹³. In contrast, most chondrules originated in a ^{16}O -poor ($\Delta^{17}\text{O} > -5\text{‰}$) gaseous reservoir at low ($<1,000$ K) ambient temperatures and higher total pressure or dust/gas ratios than CAIs^{2,14–16}. Melting of pre-existing solids accompanied by evaporation-recondensation is believed to have been the dominant process during chondrule formation^{2,14–16}.

Most CAIs show large ^{26}Mg excesses ($^{26}\text{Mg}^*$), produced by the *in situ* decay of ^{26}Al (half-life $t_{1/2} = 0.73$ Myr), corresponding to an initial $^{26}\text{Al}/^{27}\text{Al}$ ratio, $(^{26}\text{Al}/^{27}\text{Al})_0$, of $\sim 5 \times 10^{-5}$ (called 'canonical'), whereas most chondrules have smaller $^{26}\text{Mg}^*$ corresponding to $(^{26}\text{Al}/^{27}\text{Al})_0$ ratios of $\leq 1.2 \times 10^{-5}$ (ref. 11 and references therein). On the basis of these observations and the assumption that ^{26}Al was uniformly distributed in the solar nebula, it is generally inferred that CAIs formed at least 1–1.5 Myr before chondrules¹¹. This conclusion has recently been questioned¹⁷ on the basis of the new lead⁴ and magnesium¹⁷ isotopic measurements. The ^{207}Pb – ^{206}Pb ages of a group of Allende chondrules ($4,566.7 \pm 1.0$ Myr)⁴ cannot be distinguished from those of CV CAIs ($4,567.2 \pm 0.6$ Myr)². Bizzarro *et al.*¹⁷ reported high precision Mg isotope analyses of micro-drilled Allende chondrules. Data for 15 chondrules show model isochrons with initial $^{26}\text{Al}/^{27}\text{Al}$ ratios ranging from $(1.4 \pm 0.5) \times 10^{-5}$ to $(5.7 \pm 0.8) \times 10^{-5}$, systematically higher than ion probe data for Allende chondrules. At face value, these results suggest that chondrule formation began contemporaneously with the formation of CAIs, and continued for at least 1.5 Myr. We note, however, that the $(^{26}\text{Al}/^{27}\text{Al})_0$ ratios inferred from bulk Mg isotope measurements of chondrules¹⁷ may date the time for the formation of chondrule precursor materials, not the time of chondrule melting; the latter requires Mg isotope measurements of mineral separates or individual mineral grains, which have not been completed yet. In addition, spatial heterogeneity in ^{26}Al distribution in the solar nebula cannot be ruled out. On the other hand, the relative timing of CAI and chondrule formation can be resolved by studying compound objects composed of chondrule and CAI, because both constituents of such objects were affected by the same heating episode. With one exception, all CAI-chondrule compound objects consist of relict CAIs inside chondrules, suggesting that the host chondrules formed by melting of solid precursors containing pre-existing CAIs^{6–10}. The only exception is the chondrule-bearing CAI A5 from the Yamato-81020 chondrite that has been interpreted as providing evidence for chondrule formation pre-dating the formation of CAIs⁵. Here we report new observations of the mineralogy, petrography and oxygen and magnesium isotopic compositions of two chondrule-bearing CAIs (ABC and TS26) from Allende, which provide important constraints on the relative chronology of CAI and chondrule formation. Although the mineralogy and petrology of both CAIs were

described previously^{18,19}, the presence of relict chondrule material inside them remained unnoticed.

ABC is a coarse-grained, igneous, anorthite-rich (type C) CAI fragment composed of lath-shaped anorthite (99 mole% anorthite; An_{99}) and Cr-poor Al-Ti-diopside, both poikilitically enclosing spinel grains, and interstitial, åkermanite-rich (Åk_{74}) melilite

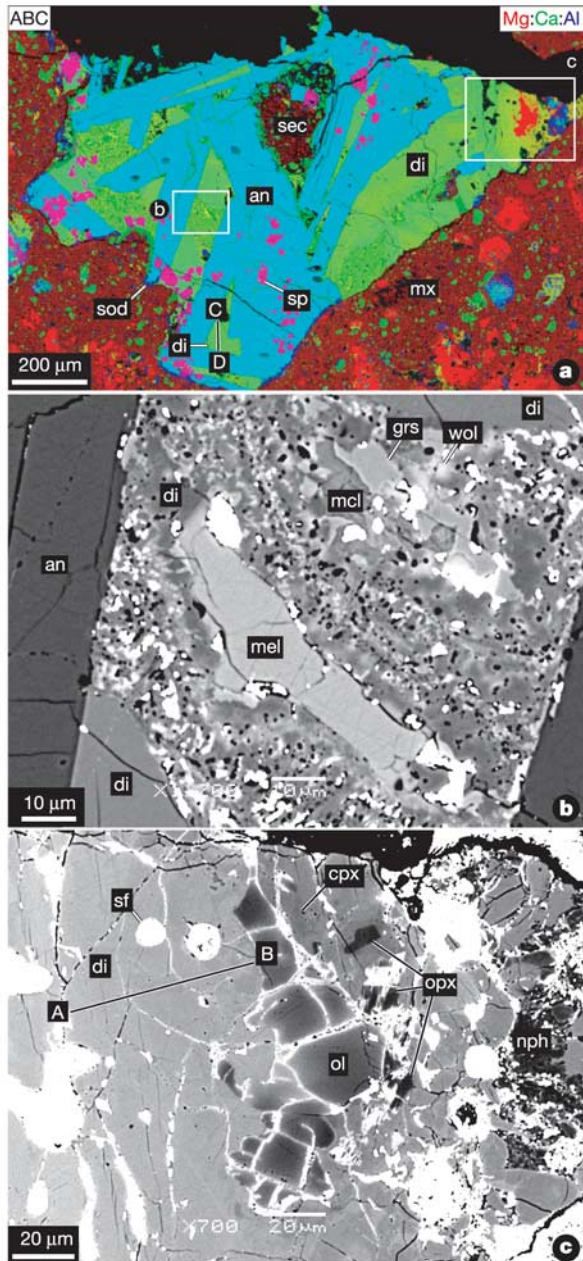


Figure 1 Type C CAI fragment ABC from Allende. **a**, Combined elemental map in Mg (red), Ca (green) and Al $K\alpha$ (blue) X-rays. **b**, **c**, Backscattered electron images. Regions outlined in **a** are shown in detail in **b** and **c**. The CAI consists of coarse-grained, anorthite laths partly replaced by sodalite and nepheline, coarse-grained Al-Ti-diopside enclosing spinel grains, and interstitial material composed of fine-grained Al-diopside and melilite replaced by grossular, monticellite and wollastonite. A relict olivine-low-Ca pyroxene fragment occurs in the right portion of the CAI in the boxed area; it is surrounded by a halo of high-Ca pyroxene. A–B (in **c**) and C–D (in **a**) indicate locations of compositional profiles shown in Supplementary Fig. 1. Abbreviations: an, anorthite; cpx, high-Ca pyroxene; di, Al-Ti-diopside; grs, grossular; mcl, monticellite; mel, melilite; nph, nepheline; ol, olivine; opx, low-Ca pyroxene; sod, sodalite; sf, Fe,Ni-sulphide; sp, spinel; wol, wollastonite.

(Fig. 1a, b; Supplementary Table 1). A coarse fragment of forsteritic olivine (5 mole% fayalite; Fa_5) intergrown with low-Ca pyroxene (1 mole% ferrosilite, 4 mole% wollastonite; Fs_1Wo_4) occurs in the CAI portion containing Cr-rich, Al-Ti-poor diopside (Fig. 1a, c; Supplementary Table 1). The olivine-pyroxene fragment is corroded by the diopside and surrounded by a halo of high-Ca pyroxene ($Fs_{0.2-0.6}Wo_{30-40}$; Fig. 1c). Olivine and low-Ca pyroxene have ^{16}O -poor compositions; spinel and Al-Ti-diopside are moderately ^{16}O -enriched, whereas Cr-spinel, Al-Ti-poor diopside, high-Ca

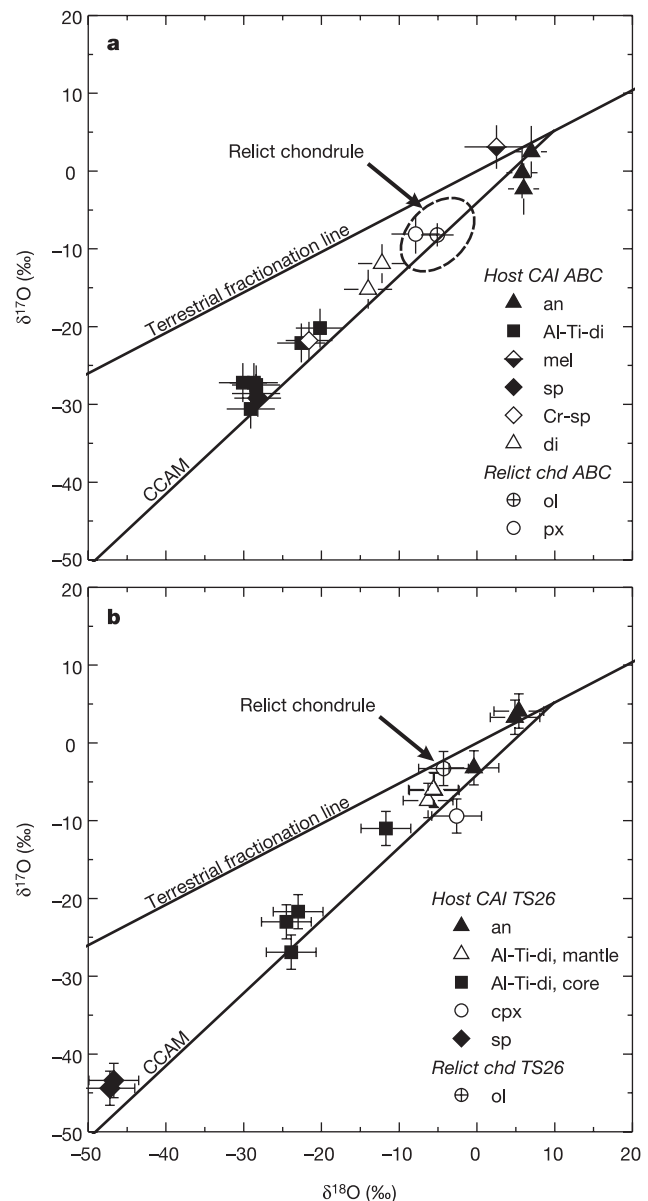


Figure 2 Oxygen isotopic compositions of chondrule-bearing CAIs from Allende. **a**, Chondrule-bearing CAI ABC. Relict olivine and low-Ca pyroxene grains have ^{16}O -poor compositions. Spinel and Al-Ti-diopside of the host CAI are ^{16}O -enriched, whereas Cr-spinel, Al-Ti-poor diopside, high-Ca pyroxene, anorthite and melilite are ^{16}O -depleted to various degrees. **b**, Chondrule-bearing CAI TS26. Relict olivine and high-Ca pyroxene halo around it have ^{16}O -poor compositions. Spinel of the CAI core is ^{16}O -rich, whereas Al-Ti-diopside and anorthite are ^{16}O -depleted to various degrees. The Al-Ti-diopside grains in the core are less ^{16}O -depleted than those in the mantle. Abbreviations: Al-Ti-di, Al-Ti-diopside; an, anorthite; chd, chondrule; di, Al-Ti-poor diopside; px, a mixture of high-Ca pyroxene (~70%) and low-Ca pyroxene (~30%); CCAM, carbonaceous chondrite anhydrous mineral line. Other abbreviations as Fig. 1. Terrestrial fractionation line and CCAM line are shown for reference. Error bars are 2σ .

pyroxene, anorthite and melilite are ^{16}O -depleted to varying degrees (Fig. 2a). The CAI shows a resolvable $^{26}\text{Mg}^*$ corresponding to a $(^{26}\text{Al}/^{27}\text{Al})_0$ ratio of $(4.7 \pm 1.4) \times 10^{-6}$ (Fig. 3).

TS26 is an irregularly-shaped type C CAI that shows a well-defined core-mantle structure (Fig. 4a; Supplementary Fig. 2), but lacks the Wark–Lovering rim layers observed around most coarse-grained CAIs from Allende¹³. It has a coarse-grained core composed of lath-shaped anorthite (An_{99}) and sector-zoned Al-Ti-diopside, both poikilitically enclosing spinel grains, and interstitial åkermanite-rich (Åk_{72}) melilite. The finer-grained mantle is composed of Al-Ti-diopside, lath-shaped anorthite, and abundant coarse grains of forsteritic olivine (Fa_{8-17}) and low-Ca pyroxene ($\text{Fs}_{1}\text{Wo}_{1-4}$). The olivine and low-Ca pyroxene grains are corroded by the diopside and surrounded by haloes of high-Ca pyroxene ($\text{Fs}_{0.4-0.6}\text{Wo}_{35-42}$; Fig. 4c). Olivine and high-Ca pyroxene have ^{16}O -poor compositions; spinel is ^{16}O -rich, whereas Al-Ti-diopside and anorthite are ^{16}O -depleted to varying degrees (Fig. 2b). The coarse Al-Ti-diopside grains in the core are less ^{16}O -depleted compared to those in the finer-grained mantle. The CAI anorthite, spinel and Al-Ti-diopside show no resolvable $^{26}\text{Mg}^*$; the inferred $(^{26}\text{Al}/^{27}\text{Al})_0$ ratio is $<1.2 \times 10^{-6}$ (Fig. 3).

The corroded appearance of olivine-pyroxene fragments in ABC and TS26 and the presence of high-Ca pyroxene haloes suggest that these grains were present inside the host CAIs during final solidification and were partly dissolved in the CAI melt. The relict origin of the olivine-pyroxene fragments is consistent with their dissolution textures and with the absence of olivine and low-Ca pyroxene in the crystallization sequence (spinel $\rightarrow$ anorthite $\rightarrow$ Al-Ti-diopside $\rightarrow$ melilite) predicted for a melt having ABC- or TS26-like bulk composition (see Supplementary Fig. 3). The coarse-grained nature of relict forsteritic olivine associated with low-Ca pyroxene and their ^{16}O -poor compositions suggest that these grains are probably fragments of ferromagnesian chondrules. Although coarse olivine grains occasionally associated with low-Ca pyroxene are also found in amoeboid olivine aggregates and in forsterite-rich accretionary rims around CAIs, these olivines and pyroxenes have characteristic ^{16}O -rich compositions^{20,21}.

Most coarse-grained igneous CAIs in Allende, including TS26 and ABC, show oxygen isotopic heterogeneity: spinel and Al-Ti-diopside are typically ^{16}O -rich ($\Delta^{17}\text{O} \approx -20\text{‰}$), whereas melilite and anorthite are ^{16}O -depleted ($\Delta^{17}\text{O}$ up to 5‰)²²⁻²⁴. This heterogeneity has recently been attributed to oxygen isotopic

exchange between an ^{16}O -poor nebular gas and initially uniformly ^{16}O -rich CAIs during incomplete melting^{23,24}. In addition, TS26 and ABC show significant ^{16}O -depletion in Al-diopside; the degree of depletion increases towards the relict chondrule fragments and the CAI peripheries. On the basis of these observations, we infer that ABC and TS26 experienced incomplete oxygen isotopic exchange and dilution with ^{16}O -poor relict chondrule materials during their

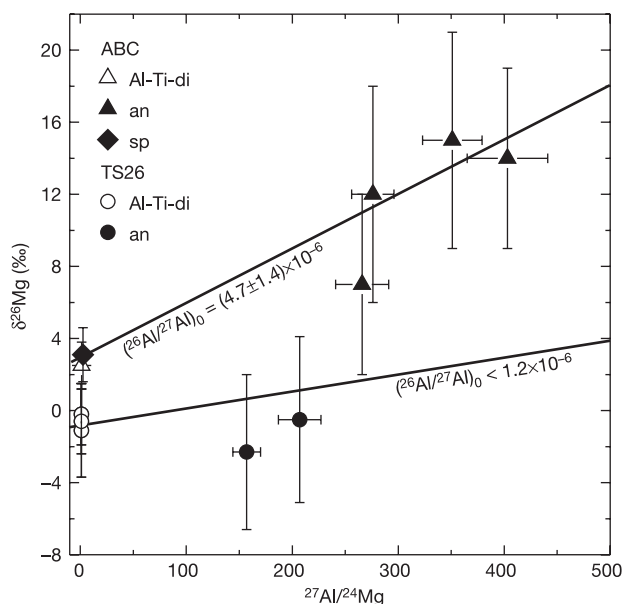


Figure 3 Al-Mg evolution diagram for the type C CAIs ABC and TS26. Error bars are 2σ .

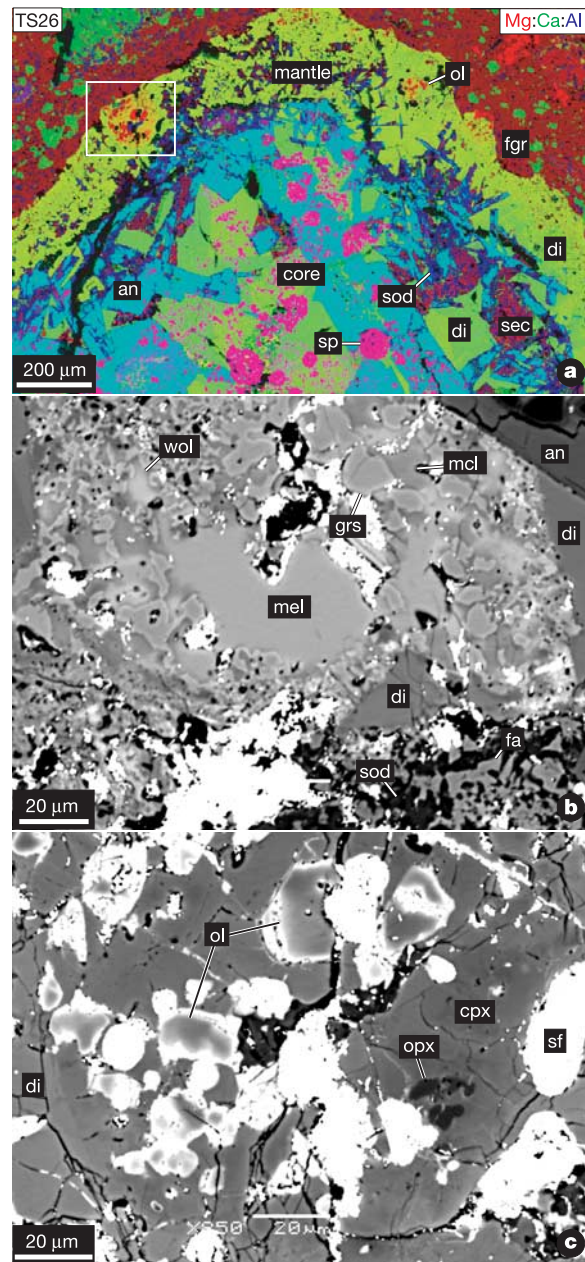


Figure 4 Type C CAI TS26 from Allende. **a**, Combined elemental map in Mg (red), Ca (green) and Al $K\alpha$ (blue) X-rays. **b**, **c**, Backscattered electron images. Region outlined in **a** is shown in detail in **c**. The entire inclusion is shown in Supplementary Fig. 2. Region shown in **b** is outlined in Supplementary Fig. 2b. The CAI has a coarse-grained core composed of anorthite laths partly replaced by sodalite, Al-Ti-diopside enclosing spinel grains, and interstitial melilite replaced by grossular, monticellite, wollastonite, sodalite and ferrous olivine. The core is surrounded by a thick Al-diopside mantle containing abundant relict fragments of olivine and low-Ca pyroxene which are surrounded by haloes of high-Ca pyroxene. The mantle is separated from the core by a discontinuous layer of Fe-Ni-sulphides. The CAI is surrounded by a fine-grained rim largely composed of ferrous olivine. Abbreviations: fa, ferrous olivine; fgr, fine-grained rim. Other abbreviations as Fig. 1.

last melting in an ^{16}O -poor gas, probably in the chondrule-forming region.

The observed differences in grain size between the core and the mantle of TS26 (see Supplementary Fig. 2a) suggests that melting was incomplete and was followed by relatively fast cooling. The absence of Wark–Lovering rim layers around TS26 could also be due to the inferred melting episode. The high abundance of relict chondrule-like material in the outer portion of TS26 (see Supplementary Fig. 2b) suggests there was a high abundance of dust in the region where melting occurred, consistent with the dusty environment inferred for chondrule formation². The low $(^{26}\text{Al}/^{27}\text{Al})_0$ ratios observed in ABC and TS26 may have recorded their late-stage re-melting during incorporation of the chondrule fragments. We note, however, that because Allende experienced thermal metamorphism that may have disturbed the ^{26}Al – ^{26}Mg systematics in CAIs and chondrules²⁵, the exact age difference between the formation of CAIs ABC and TS26 and their re-melting should be considered with caution.

The proposed multi-stage formation history of ABC and TS26 is consistent with the extended (~ 2 Myr) formation time of several other igneous CAI from CV chondrites inferred from a range of the $(^{26}\text{Al}/^{27}\text{Al})_0$ ratios within a single inclusion and petrographic observations^{26,27}. The late-stage melting and oxygen isotopic exchange of ABC and TS26 are also consistent with the recently proposed model for the global evolution of the oxygen isotope composition of the inner solar nebula gas from ^{16}O -rich to ^{16}O -poor with time^{28,29}. The fact that many CAIs show no evidence for being affected by chondrule heating suggests that the chondrule-forming events were highly localized. □

Received 17 December 2004; accepted 11 February 2005; doi:10.1038/nature03470.

- MacPherson, G. J. in *Treatise on Geochemistry* (eds Holland, H. D. & Turekian, K. K.) Vol. 1, *Meteorites, Comets and Planets* (ed. Davis, A. M.) 201–246 (Elsevier-Pergamon, Oxford, 2003).
- Desch, S. J. & Connolly, H. C. A model of the thermal processing of particles in the solar nebula shocks: Application to the cooling rates of chondrules. *Meteorit. Planet. Sci.* **37**, 183–207 (2002).
- Amelin, Y., Krot, A. N., Hutcheon, I. D. & Ulyanov, A. A. Lead isotopic ages of chondrules and calcium–aluminum-rich inclusions. *Science* **297**, 1678–1683 (2002).
- Amelin, Y., Krot, A. N. & Twelker, E. Pb isotopic age of the CB chondrite Gujba, and the duration of the chondrule formation interval. *Geochim. Cosmochim. Acta* **68**, abstr. E958 (2004).
- Itoh, S. & Yurimoto, H. Contemporaneous formation of chondrules and refractory inclusions in the early Solar System. *Nature* **423**, 728–731 (2003).
- Krot, A. N. & Keil, K. Anorthite-rich chondrules in CR and CH carbonaceous chondrites: Genetic link between Ca, Al-rich inclusions and ferromagnesian chondrules. *Meteorit. Planet. Sci.* **37**, 91–111 (2002).
- Krot, A. N., Hutcheon, I. D. & Keil, K. Anorthite-rich chondrules in the reduced CV chondrites: evidence for complex formation history and genetic links between CAIs and ferromagnesian chondrules. *Meteorit. Planet. Sci.* **37**, 155–182 (2002).
- Krot, A. N. *et al.* Ca, Al-rich inclusions, amoeboid olivine aggregates, and Al-rich chondrules from the unique carbonaceous chondrite Acfer 094: I. Mineralogy and petrology. *Geochim. Cosmochim. Acta* **68**, 2167–2184 (2004).
- Maruyama, S., Yurimoto, H. & Sueno, S. Oxygen isotope evidence regarding the formation of spinel-bearing chondrules. *Earth Planet. Sci. Lett.* **169**, 165–171 (1999).
- Maruyama, S. & Yurimoto, H. Relationships among O, Mg isotopes and the petrography of two spinel-bearing chondrules. *Geochim. Cosmochim. Acta* **67**, 3943–3957 (2003).
- McKeegan, K. D. & Davis, A. M. in *Treatise on Geochemistry* (eds Holland, H. D. & Turekian, K. K.) Vol. 1, *Meteorites, Comets and Planets* (ed. Davis, A. M.) 431–461 (Elsevier-Pergamon, Oxford, 2003).
- Grossman, L., Ebel, D. S. & Simon, S. B. Formation of refractory inclusions by evaporation of condensate precursors. *Geochim. Cosmochim. Acta* **66**, 145–161 (2002).
- Wark, D. A. & Lovering, J. F. Marker events in the early solar system: Evidence from rims on Ca–Al-rich inclusions in carbonaceous chondrites. *Proc. Lunar Planet. Sci. Conf.* **8**, 95–112 (1977).
- Galy, A., Young, E. D., Ash, R. D. & O’Nions, R. K. The formation of chondrules at high gas pressures in the solar nebula. *Science* **290**, 1751–1753 (2000).
- Alexander, C. M. O’D. & Wang, J. Iron isotopes in chondrules: Implications for the role of evaporation during chondrule formation. *Meteorit. Planet. Sci.* **36**, 419–428 (2001).
- Scott, E. R. D. & Krot, A. N. in *Treatise on Geochemistry* (eds Holland, H. D. & Turekian, K. K.) Vol. 1, *Meteorites, Comets and Planets* (ed. Davis, A. M.) 143–200 (Elsevier-Pergamon, Oxford, 2003).
- Bizzarro, M., Baker, J. A. & Haack, H. Mg isotope evidence for contemporaneous formation of chondrules and refractory inclusions. *Nature* **431**, 275–278 (2004).
- MacDougall, J. D., Kerridge, J. F. & Phinney, D. Refractory ABC. *Lunar Planet. Sci.* **12**, 643–645 (1981).
- Wark, D. A. Plagioclase-rich inclusions in carbonaceous chondrite meteorites: Liquid condensates? *Geochim. Cosmochim. Acta* **51**, 221–242 (1987).
- Krot, A. N., McKeegan, K. D., Leshin, L. A., MacPherson, G. J. & Scott, E. R. D. Existence of an ^{16}O -rich gaseous reservoir in the solar nebula. *Science* **295**, 1051–1054 (2002).
- Krot, A. N., Fagan, T. J., Yurimoto, H. & Petaev, M. I. Origin of low-Ca pyroxene in amoeboid olivine aggregates: Evidence from oxygen isotopic compositions. *Geochim. Cosmochim. Acta* (in the press).
- McKeegan, K. D. & Leshin, L. A. in *Stable Isotope Geochemistry* (eds Valley, J. W. & Cole, D. R.) 279–378 (Reviews in Mineralogy & Geochemistry, Vol. 43, Mineralogical Society of America, Washington DC, 2001).

- Yurimoto, H., Ito, M. & Nagasawa, H. Oxygen isotope exchange between refractory inclusion in Allende and solar nebula gas. *Science* **282**, 1874–1877 (1998).
- Nagashima, K., Yoshitake, M. & Yurimoto, H. in *Workshop on “Chondrites and the Protoplanetary Disk”* 153–154 (University of Hawaii at Manoa, 2004); available at (<http://www.lpi.usra.edu/meetings/chondrites2004/pdf/9072.pdf>).
- Yurimoto, H., Koike, O., Nagahara, H., Morioka, M. & Nagasawa, H. Heterogeneous distribution of Mg isotopes in anorthite single crystal from Type B CAIs in Allende meteorite. *Lunar Planet. Sci.* **31**, 1593 (2000).
- MacPherson, G. J. & Davis, A. M. A petrologic and ion microprobe study of a Vigarano Type B refractory inclusion: Evolution by multiple stages of alteration and melting. *Geochim. Cosmochim. Acta* **57**, 231–243 (1989).
- Hsu, W., Wasserburg, G. J. & Huss, G. R. High time resolution by use of the ^{26}Al chronometer in the multistage formation of a CAI. *Earth Planet. Sci. Lett.* **182**, 15–29 (2000).
- Yurimoto, H. & Kuramoto, K. Molecular cloud origin for the oxygen isotope heterogeneity in the solar system. *Science* **305**, 1763–1766 (2004).
- Krot, A. N. *et al.* Evolution of oxygen isotopic composition in the inner solar nebula *Astrophys. J.* (in the press).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements Financial support for this project was provided by NASA (A.N.K., I.D.H., G.J.M.) and Monksasho (H.Y.). We thank R. H. Hewins for comments and suggestions.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.N.K. (sasha@higp.hawaii.edu).

Controlled multiple quantum coherences of nuclear spins in a nanometre-scale device

Go Yusa¹, Koji Muraki¹, Kei Takashina¹, Katsushi Hashimoto^{2*} & Yoshiro Hirayama^{1,2}

¹NTT Basic Research Laboratories, NTT Corporation, 3-1 Morinosato-Wakamiya, Atsugi 243-0198, Japan

²SORST Program, Japan Science and Technology Agency (JST), 4-1-8 Honmachi, Kawaguchi, Saitama 331-0012, Japan

* Present address: Institute of Applied Physics, Hamburg University, Jungiusstraße 11, D-20355 Hamburg, Germany

The analytical technique of nuclear magnetic resonance (NMR)^{1,2} is based on coherent quantum mechanical superposition of nuclear spin states. Recently, NMR has received considerable renewed interest in the context of quantum computation and information processing^{3–11}, which require controlled coherent qubit operations. However, standard NMR is not suitable for the implementation of realistic scalable devices, which would require all-electrical control and the means to detect microscopic quantities of coherent nuclear spins. Here we present a self-contained NMR semiconductor device that can control nuclear spins in a nanometre-scale region. Our approach enables the direct detection of (otherwise invisible) multiple quantum coherences between levels separated by more than one quantum of spin angular momentum. This microscopic high sensitivity NMR technique is especially suitable for probing materials whose nuclei contain multiple spin levels, and may form the basis of a versatile multiple qubit device.

Nuclei often possess total spin I greater than a half. Under static magnetic field B_0 , therefore, $2I + 1$ states $|m\rangle$ equally spaced in energy by the Zeeman energy $\hbar\omega_0$ are formed according to the Zeeman effect (Fig. 1e). Here, $\hbar$ is the reduced Planck’s constant such that ω_0 would be the resonant angular frequency of NMR between any pair of adjacent states. After appropriate polarization,

pulsed radio frequency (r.f.) electromagnetic field can be applied to achieve a coherent superposition $|\Psi\rangle = \sum_{m=-I}^I a_m |m\rangle$ of nuclear spins with arbitrary complex amplitude a_m by controlling two out of three readily tunable parameters^{1,12}; the time length τ_P of the r.f. pulse, the intensity B_1 of the magnetic component of the r.f. field and the offset $\hbar(\omega_R - \omega)$ between the energies of the r.f. field $\hbar\omega$ and the resonant transition $\hbar\omega_R$. Therefore, nuclei forming multiple spin levels have a large variety of possible coherent combinations. However, multiple quantum coherence is not directly detectable by conventional NMR as it does not produce transverse magnetization $\mathbf{M}_{x,y}$ (Fig. 1a), which conventional NMR detects. Our objective here is to produce a method that does not rely upon electromagnetic coupling of $\mathbf{M}_{x,y}$ to a pick-up coil^{13–16}, and to control and detect not only the single but also the multiple quantum coherences of nuclear spins in a nanoscale region by all-electrical means.

We fabricated a monolithic semiconductor device integrated with a point contact channel, which is a narrow constriction of a two-dimensional electron gas, and an antenna gate for locally irradiating the channel with an r.f. field (Fig. 1b). In this structure, we can selectively polarize nuclear spins in this nanometre-scale point contact region, while keeping those elsewhere in thermal equilibrium¹⁷ (see Methods). With polarization followed by r.f. pulse irradiation, the resistance of the point contact falls, and then gradually increases again over a time period of the order of $\sim 10^2$ – 10^3 s and reaches the steady state value at polarization. We therefore interpret ΔR , the change in resistance directly before and after the pulse, to have a direct relationship with M_z , which is the longitudinal magnetization induced by the altered population of nuclear spin states. This in turn corresponds to the projection of the probability amplitude of the nuclear spin state $|\Psi\rangle$ onto the z axis (Fig. 1a).

The channel consists of ⁶⁹Ga, ⁷¹Ga and ⁷⁵As, each having total spin $I = 3/2$. Thus, for each nuclide under B_0 (Fig. 1a), four equally spaced energy states $|m\rangle = |3/2\rangle, |1/2\rangle, |-1/2\rangle, |-3/2\rangle$ are formed (Fig. 1e, orange arrows). However, the nuclei experience the effects of an electric field gradient due to their host crystal, which shift the energy states through the electric quadrupolar interaction by energies $+\Delta$ and $-\Delta$ for $|\pm 3/2\rangle$ and $|\pm 1/2\rangle$, respectively^{1,12,18,19}. The energy difference between adjacent states is, therefore, shifted by $\pm 2\Delta$ from $\hbar\omega_0$, allowing three possible resonances, at $\hbar\omega_0 - 2\Delta$, $\hbar\omega_0$ and $\hbar\omega_0 + 2\Delta$ (Fig. 1e, red arrows). According to energy and angular momentum conservation rules, several transitions are possible. For convenience of the later discussion, we label possible transitions by superscripts, for example, ⁷⁵As^{IV/2} or ⁶⁹Ga^{IV/2}, in which numerators, I, II, ..., VI specify transitions and denominators 1, 2 and 3 represent the number of photons n , the second quantization of the r.f. field, involved in the transition.

Spectra of ΔR as a function of τ_P show three oscillations at low B_1 (Fig. 2a). These oscillations can be well fitted by damped oscillations¹² $\propto 1 - \cos(\Omega_R \tau_P) \exp(-\tau_P/T_2)$ with oscillation frequency Ω_R and decoherence time T_2 , showing that ΔR is indeed linearly related to the projection of the probability amplitude of $|\Psi\rangle$ onto the z axis and to the longitudinal magnetization M_z . In a geometrical representation^{1,12,20} of superpositional states of a two-level system (the Bloch sphere), the oscillations at $\hbar\omega_0$ (denoted by ⁷⁵As^{II} in Fig. 2d) correspond to the rotation of a state $|\Psi^{II}\rangle$ whose path describes a great circle passing through localized states $|1/2\rangle$ and $|-1/2\rangle$ periodically as the angle $\Omega_R \tau_P$ is swept. Therefore, the oscillations at $\hbar\omega_0$ represent a single quantum coherence between levels separated by a single quantum of angular momentum ($\Delta m = 1$). Similar oscillations seen in the lower- and higher-energy peaks of the spectra (denoted respectively by ⁷⁵As^I and ⁷⁵As^{III} in Fig. 2d) are also due to single quantum coherence, their energy shift being $\pm 2\Delta$ from $\hbar\omega_0$.

When B_1 is increased, two slow oscillations (denoted by ⁷⁵As^{IV/2,V/2} in Fig. 2e) are observed in the spectra (Fig. 2b) between the three single quantum oscillations. The peak energies of ⁷⁵As^{IV/2,V/2} are

shifted from $\hbar\omega_0$ by $\pm \Delta$. When the peak energies of ⁷⁵As^{IV/2,V/2} are doubled ($= 2(\hbar\omega_0 \pm \Delta)$), the resulting values are exactly the energy separations between $|-3/2\rangle$ and $|1/2\rangle$ and between $|-1/2\rangle$ and $|3/2\rangle$ (Fig. 1e). These can be assigned to double quantum coherent oscillations between two levels separated by two quanta of angular momentum ($\Delta m = 2$) driven by two-photon transitions (IV/2 and V/2 in Fig. 1e). By increasing B_1 further, a new peak (denoted by ⁷⁵As^{VI/3} in Fig. 2f) appears at $\hbar\omega_0$ (Fig. 2c). The energy difference between $|-3/2\rangle$ and $|3/2\rangle$ is equal to $3\hbar\omega_0$ regardless of the quadrupolar interaction (Fig. 1e). Therefore, the narrow peak (VI/3) can be assigned to triple quantum coherence between $|-3/2\rangle$ and

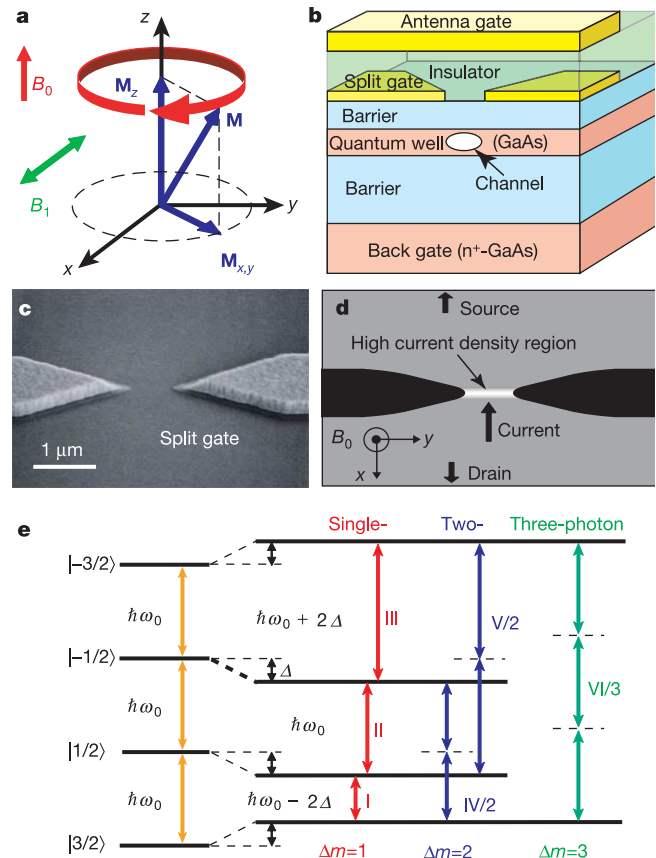


Figure 1 Schematic diagrams illustrating the main features of our device and the experimental system. **a**, Schematic illustration of the magnetization vector $\mathbf{M}$ of an ensemble of nuclear spins under a static magnetic field B_0 and an r.f. field. (B_1 is the magnetic field component of the r.f. field.) The transverse component $\mathbf{M}_{x,y}$ is precessing in the x - y plane (Larmor precession), while the longitudinal component M_z is static along B_0 . **b**, Schematic illustration of a cross-sectional view of the device. The structure contains a 20-nm GaAs quantum well with AlGaAs barrier layers and an n^+ -GaAs substrate that functions as a back-gate electrode to control electron density. A pair of Schottky split gates defines the point contact channel indicated by a white ellipse. The antenna gate is electrically isolated from the split gates by an insulator. **c**, Scanning electron microscope image of the split gates beneath the antenna gate. The gap between the gates is 600 nm, which is narrow enough to observe quantized conductance steps of $2e^2/h$ at $B_0 = 0$, where h is Planck's constant and e is the elementary charge. **d**, Schematic illustration of the plane view of the point contact channel. By applying negative voltage to the split gate, the depletion regions (shown in black) are formed. By passing current (~ 7 nA) between the source and drain at Landau level filling factor $\nu = 2/3$ (refs 26–28), nuclear spins are polarized in the point contact region, where the current density is locally high. The resistance across the point contact is measured by a four-terminal method in constant-current mode to achieve a steady state of nuclear spin polarization¹⁷. **e**, Schematic energy level diagram of nuclear spin states for $I = 3/2$ with (right) and without (left) electric quadrupolar interactions. See main text for nomenclature details.

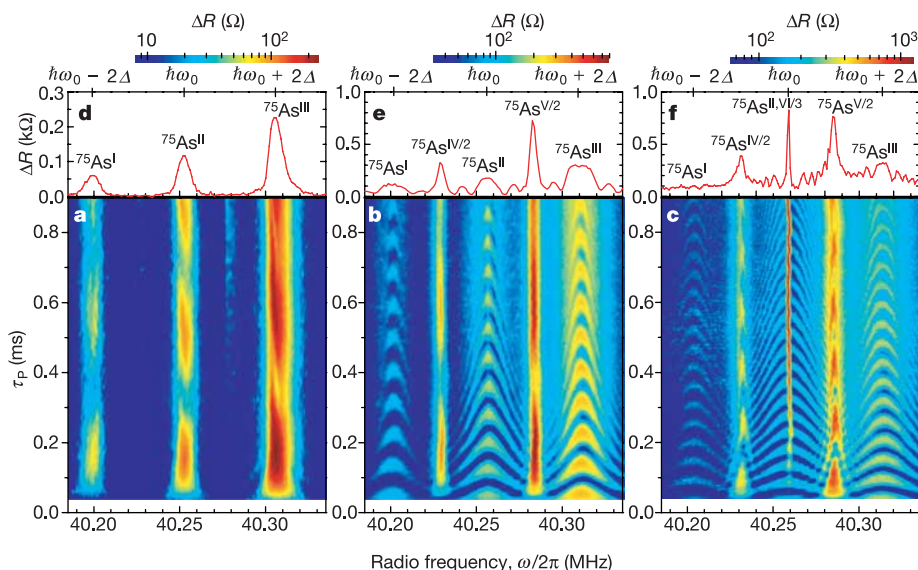


Figure 2 Spectra of ΔR for ^{75}As at three different intensities B_1 , which is proportional to the square root of the output power $P_{\text{r.f.}}$ from an r.f. generator. The measurement was performed at $B_0 = 5.5$ T at 0.1 K. ΔR is colour coded on a logarithmic scale for clarity. (The resolution is not equal for the three panels.) **a**, Colour plot of ΔR as a function of τ_p at

$P_{\text{r.f.}} = 0$ dBm. **b**, 13 dBm. **c**, 20 dBm. **d**, Spectrum at $\tau_p = 0.12$ ms taken from **a**. **e**, Spectrum at $\tau_p = 0.126$ ms taken from **b**. **f**, Spectrum at $\tau_p = 0.426$ ms taken from **c**.

$|3/2\rangle$, involving three photons with energy $\hbar\omega_0$. Single and triple quantum coherences are produced and detected by a single r.f. pulse.

All six coherences are observed for all three nuclides (^{75}As , ^{69}Ga and ^{71}Ga). Whereas the observation of clear oscillations due to the three-photon transition is hampered by r.f. power limitations for $^{75}\text{As}^{VI/3}$, such a high power regime is readily accessible for ^{69}Ga , which has a larger gyromagnetic ratio^{1,12,18,19}. ΔR of ^{69}Ga as a function of τ_p at $\hbar\omega_0$ reveals slow oscillations of $^{69}\text{Ga}^{VI/3}$ superimposed on the fast oscillations of $^{69}\text{Ga}^{II}$ (Fig. 3a). These two oscillations clearly show different dependence on B_1 in their oscillation frequencies Ω_R (Fig. 3b). At low B_1 , only the fast

oscillations of $^{69}\text{Ga}^{II}$ are visible (line A in Fig. 3a). With increasing B_1 , the slow oscillations of $^{69}\text{Ga}^{VI/3}$ start to appear (line B), and its frequency rapidly increases in a manner expected for a multi-photon process^{21,22}.

As coherent oscillations are observed at any resonant energy $\hbar\omega_R$, coherent superpositions $|\Psi\rangle$ of any two states, which lie on the great circle of the Bloch sphere, are achieved. As $\hbar\omega$ shifts away from $\hbar\omega_R$, the oscillation frequency Ω increases while the oscillation amplitude decreases, resulting in off-resonance tails (Fig. 2). This behaviour corresponds to the trajectory of $|\Psi\rangle$ for constant $\hbar\omega$, moving from the great circle on the Bloch sphere, shrinking towards convergence

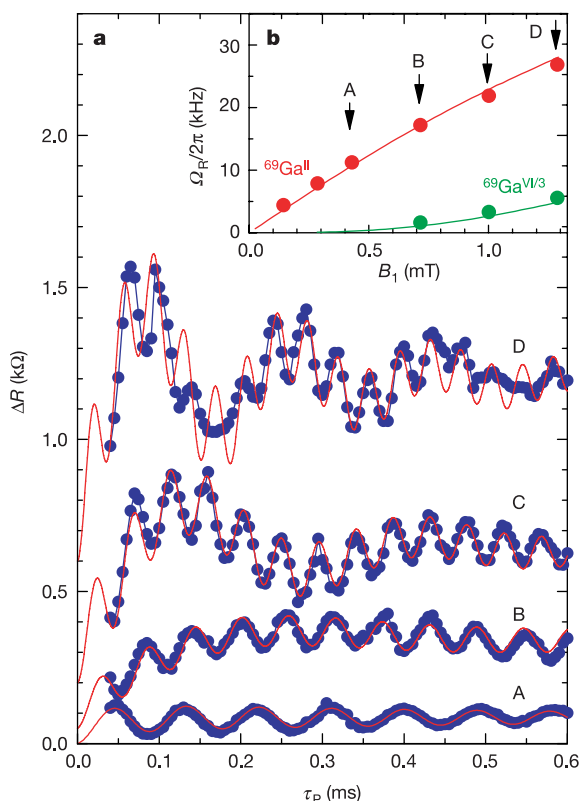


Figure 3 Behaviour of ΔR for ^{69}Ga at $\omega_0/2\pi$. **a**, ΔR at $\omega_0/2\pi = 64.7141$ MHz for ^{69}Ga as a function of τ_p at several different values of B_1 . Line A: $P_{\text{r.f.}} = 9.54$ dBm, B: 14 dBm, C: 17 dBm, D: 19 dBm. The measurement was performed at $B_0 = 6.3$ T at 75 mK. The data have been offset for clarity. Red lines are phenomenologically fitted by two oscillatory components with two different Ω_{RS} and T_2 s by $|i\rangle > \propto 1 - \cos(\Omega_{\text{R}}\tau_p) \exp(-\tau_p/T_2)$. The T_2 of the fast oscillations of $^{69}\text{Ga}^{II}$ is between ~ 0.5 and 1.5 ms, while that of $^{69}\text{Ga}^{VI/3}$ is ~ 0.2 ms. In line D, at the first and third minima of the slow oscillations of $^{69}\text{Ga}^{VI/3}$, the amplitude of $^{69}\text{Ga}^{II}$ is clearly smaller and the oscillation frequency of $^{69}\text{Ga}^{II}$ is faster than the fit. Also, similar discrepancy from the simple superpositional oscillations is observed at $\hbar\omega_0$ for ^{75}As . **b**, Oscillation frequency $\Omega_{\text{R}}/2\pi$ as a function of B_1 for $^{69}\text{Ga}^{II,VI/3}$. Lines show the theoretical behaviour obtained by taking Fourier transforms of the numerical calculations determined in the main text. The quadrupolar frequency (15.2 kHz) and the gyromagnetic ratio ($1.17 \times 10^7 \text{ s}^{-1} \text{ T}^{-1}$) used in the calculation are estimated from the experiment.

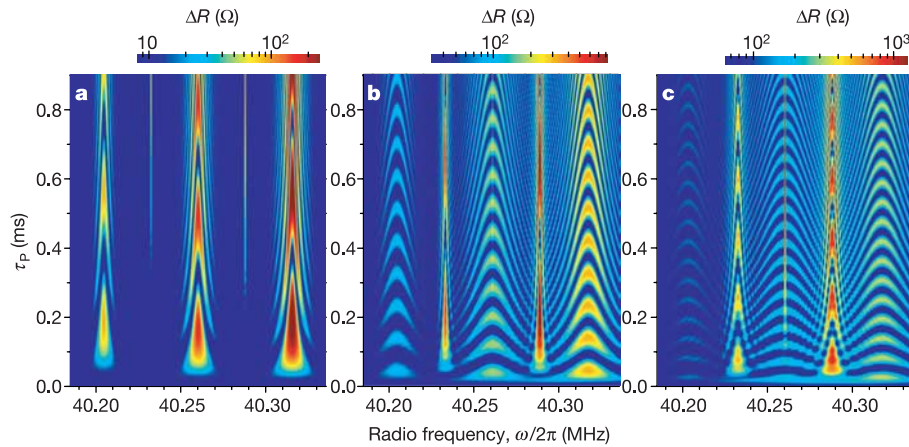


Figure 4 Colour plots showing calculated values of ΔR . The exact quantum mechanical evolution of ^{75}As is numerically simulated using the rotating frame approximation^{1,10,12,23}. The nuclear spin hamiltonian used here is $H = H_z + H_Q + \gamma \hbar B_1 \cos(\omega t) I_x$. Here $H_z = -\gamma \hbar B_0 I_z$ is the Zeeman term, where γ is the gyromagnetic ratio of each nuclide; $H_Q = \Delta/3[3I_z^2 - I(I+1)]$ is the quadrupolar term, and $\gamma \hbar B_1 \cos(\omega t) I_x$ is the magnetic field term of the r.f. H is transformed into H_{rot} in the rotating frame, and used to calculate

the expectation value of M_z (refs 10, 23). Relaxation processes are not taken into account. $\Delta/h = 26.9 \text{ kHz}$ and $\gamma = 7.32 \times 10^6 \text{ s}^{-1} \text{ T}^{-1}$ are estimated from Fig. 2. The initial population ρ_0 and the conversion coefficient between the change in the resistance and the calculated evolution of M_z are determined by fitting the spectrum in Fig. 2c at $\tau_p = 82 \mu\text{s}$, which is short compared with $T_2 \approx 1 \text{ ms}$. **a**, $P_{\text{r.f.}} = 0 \text{ dBm}$ ($B_1 = 0.2 \text{ mT}$); **b**, 13 dBm (0.85 mT); **c**, 20 dBm (1.4 mT).

at one localized state $|m\rangle$. This demonstrates that coherent superpositions of any two states with any order Δm of quantum coherence can be achieved with any arbitrary complex amplitude by a single r.f. pulse at constant B_1 , controlled by τ_p and offset energy $\hbar(\omega_R - \omega)$. (See Methods.)

We have simulated the quantum mechanical evolution and thereby the expectation value of the longitudinal magnetization M_z of an ensemble of spin $I = 3/2$ systems using the rotating-frame approximation^{1,10,12,23} (Fig. 4). The calculation, in which the only input parameter is the r.f. power B_1 , can reproduce the overall features of the data (Fig. 2) remarkably well, including not only the frequencies of all observed oscillations but also the behaviour of off-resonance tails. Their widths for higher Δm are scaled down by the photon number n and are therefore narrower. The oscillation amplitudes for higher Δm are larger than those for $\Delta m = 1$, reflecting greater change in the angular momentum Δm . The calculated B_1 dependence of the oscillation frequency agrees well with the data (Fig. 3b). These excellent agreements, both qualitative and quantitative, justify our interpretation in terms of multi-photon processes. Furthermore, fitting the relative oscillation amplitudes of different transitions allows us to determine the initial polarization before an r.f. pulse $\rho_0 = \{\rho_{|3/2\rangle}, \rho_{|1/2\rangle}, \rho_{|-1/2\rangle}, \rho_{|-3/2\rangle}\}$ to be $\{0, 0.1, 0.28, 0.62\}$, which is surprisingly far from equilibrium (an apparent spin temperature of -3 mK). Moreover, the fact that a single conversion coefficient of 898Ω per nuclear magneton can not only fit all data for different B_1 but also for different Δm clearly shows the linear relationship between M_z and ΔR , thus confirming our detection mechanism.

At resonant frequencies with only one coherence visible, the temporal decay of the oscillations can be fitted well by introducing a phenomenological decoherence time T_2 . At $\hbar\omega_0$, where the single and triple coherences coincide, the data can be fitted by assuming two independent T_2 s for two coherences when the r.f. power is not too high (lines A, B and C in Fig. 3a). At the highest power, however, deviation from such a simple model becomes discernible (line D). We confirm that numerical calculation, which incorporates the interference between different coherences within the same nucleus, cannot reconcile this deviation. This in turn leads us to speculate that additional effects, such as coupling between different nuclei, start to play a role.

As we have demonstrated, three single, two double and one triple quantum coherences for one nuclide, 18 in total for three nuclides,

are completely controlled by all-electrical means. One of the simplest applications of our device would be for quantum search engines^{9,10,23,24} using the Grover algorithm²⁵, which can be performed using multi-level nuclear spins. Our device is not in itself a scalable multi-qubit system, but if different coherences could be entangled—for example, by controlling coupled coherence between different nuclei as suggested above—it would allow quantum entanglement²⁰ of nuclear spins to express superpositions of different binary strings. The device is also likely to find application as a quantum-information storage device for quantum mechanical states of electrons¹¹, as its decoherence time is orders of magnitude longer than any other likely candidate. Using conduction electrons as a mediator, such devices could be integrated into solid-state quantum circuits incorporating multilevel nuclear spin systems. □

Methods

Mechanisms for polarization and detection

Under special conditions of electron density and B_0 in the fractional quantum Hall regime^{26–28}, two states with different electron spin polarizations become nearly degenerate in energy. When sufficiently large current is applied in such cases, nuclear spins in contact with the two-dimensional electron system are polarized as a result of flip-flop scattering through the contact hyperfine interaction. The hyperfine field produced by the polarized nuclei modifies the Zeeman energy of electrons and hence the back scattering probability, as manifested by a change in the resistance²⁷. In our device, this occurs only when a certain negative voltage ($\leq -0.3 \text{ V}$) is applied to the split gates, and a current above a threshold is passed. Owing to the narrow constriction (Fig. 1b–d), the local current density exceeds the threshold only in the point contact region, thereby locally producing nuclear polarization far from equilibrium. For detecting the nuclear polarization after the manipulation with the r.f. pulse, the same condition (in terms of the electron density, B_0 and the split-gate voltage) is used, as it also ensures high sensitivity of the resistance to the nuclear spin polarization. If we estimate the dimensions of the nanoscale point contact region (Fig. 1b) to be $\sim 200 \times 200 \text{ nm}^2$ (area) and $\sim 10 \text{ nm}$ (thickness), the number of nuclei in the resulting volume is of the order of $\sim 10^8$ or less, several orders of magnitude smaller than the detection limit of conventional NMR ($10^{11}–10^{13}$).

Towards coherence over all four levels

In this Letter we mainly discuss multiple coherences between two spin states out of four with a single r.f. frequency. Controlling coherence over all four levels requires an r.f. pulse with three different energies, which could be readily achieved using three r.f. generators and a conventional r.f. combiner.

Received 5 November 2004; accepted 14 February 2005; doi:10.1038/nature03456.

- Ernst, R. R., Bodenhausen, G. & Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Oxford Univ. Press, Oxford, 1987).
- Liang, Z. P. & Lauterbur, P. C. *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective* (IEEE Press, Piscataway, 1999).

3. DiVincenzo, D. P. Quantum computation. *Science* **270**, 255–261 (1995).
4. Chuang, I. L., Vandersypen, L. M. K., Zhou, X., Leung, D. W. & Lloyd, S. Experimental realization of a quantum algorithm. *Nature* **393**, 143–146 (1998).
5. Kane, B. E. A silicon-based nuclear spin quantum computer. *Nature* **393**, 133–137 (1998).
6. Nakamura, Y., Pashkin, Yu. A. & Tsai, J. S. Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature* **398**, 786–788 (1999).
7. Mooij, J. E. *et al.* Josephson persistent-current qubit. *Science* **285**, 1036–1039 (1999).
8. Vandersypen, L. M. K. *et al.* Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance. *Nature* **414**, 883–887 (2001).
9. Leuenberger, M. N. & Loss, D. Quantum computing in molecular magnets. *Nature* **410**, 789–793 (2001).
10. Leuenberger, M. N., Loss, D., Poggio, M. & Awschalom, D. D. Quantum information processing with large nuclear spins in GaAs semiconductors. *Phys. Rev. Lett.* **89**, 207601 (2002).
11. Taylor, J. M., Marcus, C. M. & Lukin, M. D. Long-lived memory for mesoscopic quantum bits. *Phys. Rev. Lett.* **90**, 206803 (2003).
12. Levitt, M. H. *Spin Dynamics* (Wiley, New York, 2002).
13. Wald, K. R., Kouwenhoven, L. P., McEuen, P. L., van der Vaart, N. C. & Foxon, C. T. Local dynamic nuclear polarization using quantum point contacts. *Phys. Rev. Lett.* **73**, 1011–1014 (1994).
14. Gammon, D. *et al.* Nuclear spectroscopy in single quantum dots: nanoscopic Raman scattering and nuclear magnetic resonance. *Science* **277**, 85–88 (1997).
15. Kikkawa, J. M. & Awschalom, D. D. All-optical magnetic resonance in semiconductors. *Science* **287**, 473–476 (2000).
16. Machida, T., Yamazaki, T., Ikushima, K. & Komiyama, S. Coherent control of nuclear-spin system in a quantum Hall device. *Appl. Phys. Lett.* **82**, 409–411 (2003).
17. Yusa, G., Hashimoto, K., Muraki, K., Saku, T. & Hirayama, Y. Self-sustaining resistance oscillations: Electron-nuclear spin coupling in mesoscopic quantum Hall devices. *Phys. Rev. B* **69**, 161–302 (2004).
18. Salis, G., Awschalom, D. D., Ohno, Y. & Ohno, H. Origin of enhanced dynamic nuclear polarization and all-optical nuclear magnetic resonance in GaAs quantum wells. *Phys. Rev. B* **64**, 195304 (2001).
19. Eickhoff, M., Lenzman, B., Flinn, G. & Suter, D. Coupling mechanisms for optically induced NMR in GaAs quantum wells. *Phys. Rev. B* **65**, 125301 (2002).
20. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, Cambridge, 2003).
21. Cohen-Tannoudji, C., Dupont-Roc, J. & Grynberg, G. *Atom-photon Interactions: Basic Processes and Applications* 488–489 (Wiley, New York, 1998).
22. Wokaun, A. & Ernst, R. R. Selective excitation and detection in multilevel spin systems: Application of single transition operators. *J. Chem. Phys.* **67**, 1752–1758 (1977).
23. Leuenberger, M. N. & Loss, D. Grover algorithm for large nuclear spins in semiconductors. *Phys. Rev. B* **68**, 165317 (2003).
24. Ahn, J., Weinacht, T. C. & Bucksbaum, P. H. Information storage and retrieval through quantum phase. *Science* **287**, 463–465 (2000).
25. Grover, L. K. Quantum computers can search arbitrarily large databases by a single query. *Phys. Rev. Lett.* **79**, 4709–4712 (1997).
26. Kronmüller, S. *et al.* New resistance maxima in the fractional quantum Hall effect regime. *Phys. Rev. Lett.* **81**, 2526–2529 (1998).
27. Hashimoto, K., Muraki, K., Saku, T. & Hirayama, Y. Electrically controlled nuclear spin polarization and relaxation by quantum-Hall states. *Phys. Rev. Lett.* **88**, 176601 (2002).
28. Smet, J. H. *et al.* Gate-voltage control of spin interactions between electrons and nuclei in a semiconductor. *Nature* **415**, 281–286 (2002).

Acknowledgements The authors are grateful to T. Fujisawa, Y. Tokura, S. Sasaki, K. Semba, S. Saito, K. Ono, S. Tarucha, T. Machida, T. Ota and N. Kumada for discussions.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.Y. (yusa@NTTBRL.jp) or K.M. (muraki@will.brl.ntt.com.jp).

Seasonal prediction of hurricane activity reaching the coast of the United States

Mark A. Saunders & Adam S. Lea

Benfield Hazard Research Centre, Department of Space and Climate Physics, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK

Much of the property damage from natural hazards in the United States is caused by landfalling hurricanes^{1–3}—strong tropical cyclones that reach the coast. For the southeastern Atlantic coast of the US, a statistical method for forecasting the occurrence of landfalling hurricanes for the season ahead has been reported⁴, but the physical mechanisms linking the predictor variables to the frequency of hurricanes remain unclear. Here we

present a statistical model that uses July wind anomalies between 1950 and 2003 to predict with significant and useful skill the wind energy of US landfalling hurricanes for the following main hurricane season (August to October). We have identified six regions over North America and over the east Pacific and North Atlantic oceans where July wind anomalies, averaged between heights of 925 and 400 mbar, exhibit a stationary and significant link to the energy of landfalling hurricanes during the subsequent hurricane season. The wind anomalies in these regions are indicative of atmospheric circulation patterns that either favour or hinder evolving hurricanes from reaching US shores.

The North Atlantic hurricane season extends from 1 June to 30 November. However, 86% of US hurricane strikes and 96% of US intense (major) hurricane strikes in 1950–2003 occurred after 1 August⁵. The large year-to-year variability in the number of hurricanes making US landfall (range zero to six since 1950) means that skilful seasonal forecasts of activity would benefit a spectrum of decision makers by reducing risk and uncertainty. Seasonal hurricane forecasting was pioneered⁶ in the mid-1980s for the North Atlantic. Assessments of seasonal forecast skill for US landfalling hurricane activity have been made for the southeast US from 1 August⁴ and for the whole US in September⁷. Work has also examined the probability of US hurricane landfall and damage as a function of the sign and strength of the El Niño/Southern Oscillation (ENSO)^{8–11}. However, with one exception⁴, none of these studies has claimed skill that is significant, robust and high enough to be practically useful. For example, the hindcast correlation skill for predicting the number of US landfalling hurricanes in September 1950–2000 is only ~0.38 (ref. 7).

Seasonal US landfalling hurricane activity is referenced usually in terms of the numbers of tropical storms, hurricanes or intense hurricanes making US landfall. We introduce the National Oceanic and Atmospheric Administration's Accumulated Cyclone Energy (ACE) index¹² as a single more appropriate measure for categorizing 'seasonal US landfalling hurricane activity'. We call this the US ACE index, and define it as the sum of the squares of hourly maximum sustained wind speeds (in units of knots) from all tropical cyclones over the US mainland (including those that have changed to extra-tropical) that have winds of at least tropical storm strength. This value is then reduced by a factor of 6 for compatibility with the ACE index at sea, which is computed from wind values every 6 h (ref. 12). The US ACE index is effectively a wind energy index indicating the cumulative wind energy from all US-striking tropical storms, hurricanes and intense hurricanes occurring during a given season. We compute the US ACE index using the maximum sustained wind speed data from the US National Hurricane Center's North Atlantic hurricane database¹³.

Our analysis also uses monthly wind data averaged between 925 and 400 mbar (about 750 to 7,000 m above sea level) from the National Center for Environmental Prediction/National Center for Atmospheric Research reanalysis¹⁴ during 1950–2003. The motion of hurricanes is determined by height-averaged winds between these levels^{15,16}. US hurricane total (economic) and insured loss data are obtained from ref. 3 (updated economic losses through to the end of 2003 are provided by C. W. Landsea) and from ref. 17 (insured losses for 2000–03 are provided by the US Property Claims Service) respectively for the period 1950–2003. All correlation coefficients (r_{rank}) refer to the Spearman rank correlation coefficient. All statistical significances are two-sided values corrected for serial autocorrelation^{18,19}.

The US ACE index 1950–2003 is linked significantly to tropospheric height-averaged wind anomalies occurring over North America, the east Pacific and the North Atlantic both before (July) and during the main Atlantic hurricane season. August, September and October are the main months for landfalling hurricane wind energy, with 82% of the annual US ACE index occurring therein. Figure 1 displays for July (Fig. 1a) and

August–October (Fig. 1b) the direction, magnitude and significance of the composite difference in wind anomalies averaged between 925 mbar and 400 mbar for those years when the US ACE index is in its upper and lower quartiles 1950–2003. The vector winds in Fig. 1 are associated with an above-median seasonal US ACE index; wind anomalies of the opposite sign are associated with a below-median seasonal US ACE index.

The composite difference in July height-averaged winds shows several areas of significance. The white boxed regions in Fig. 1a denote six areas (u_1 , u_2 , u_3 , u_4 , u_5 and v_1 , where u and v refer respectively to zonal (east–west) and meridional (north–south) winds) where the link between the July 925–400 mbar height-averaged wind and the US ACE index 1950–2003 is significant and stationary. Statistical significance is defined as a correlation P -value of <0.1 after correction for serial autocorrelation. Temporal stability is defined as showing statistical significance over each sub-period 1950–76 and 1977–2003. A prominent feature in Fig. 1a is the anticyclonic flow anomaly associated with a strengthened and northward displaced Bermuda high pressure area. When established in July, this feature tends to persist and ridge westwards during

August and September to lie over the US and Canadian eastern seaboard. This leads to on-shore wind anomalies along the US East Coast during August–October (Fig. 1b). As these anomalies will influence the steering of hurricanes, their presence favours an above-median US ACE index year.

Further prominent features in Fig. 1a that favour a subsequent above-median US ACE index are anticyclonic flow associated with increased surface pressure over the Rocky Mountains, and oppositely directed zonal flow anomalies (u_1 and u_2) in the tropical east Pacific. These latter are linked to anomalous zonal gradients in sea surface temperature (SST) caused in part by ENSO. These additional features combine to produce and sustain a cyclonic flow anomaly centred over the Gulf of Mexico during August–October (Fig. 1b). This cyclonic flow strengthens on-shore wind anomalies over Florida and the eastern US Gulf Coast, thereby favouring an above-median US ACE index. July height-averaged wind anomalies opposite to those in Fig. 1a hinder subsequent US hurricane landfall, and herald a below-median US ACE index year.

The identification of tropospheric height-averaged wind anomalies linked significantly and stably to the upcoming US ACE index allows the seasonal predictability of the latter to be assessed. Predictability is computed using cross-validated hindcasts with block elimination^{20,21}. This procedure is applied to linear regression models having a single wind index predictor that satisfy the validity assumptions for ordinary least squares linear regression^{22,23}. These models are distinguished by the regions contributing to the July height-averaged wind index. Model 1 uses a wind index defined as $u_1 - u_2 + u_3 - u_4 + u_5 - v_1$ with contributions from all six regions. Here each u and v component refers to the area-averaged wind anomaly (Fig. 1a) after normalization by standard deviation. Model 2 uses $u_1 - u_2 - v_1$ as the wind index with contributions from the east Pacific and North America regions only. Model 3 uses $u_3 - u_4 + u_5$ as the index with contributions from the North Atlantic regions only. Skill from these three July 925–400 mbar wind index models is compared to that achievable with perfect *a priori* knowledge of North Atlantic ACE activity

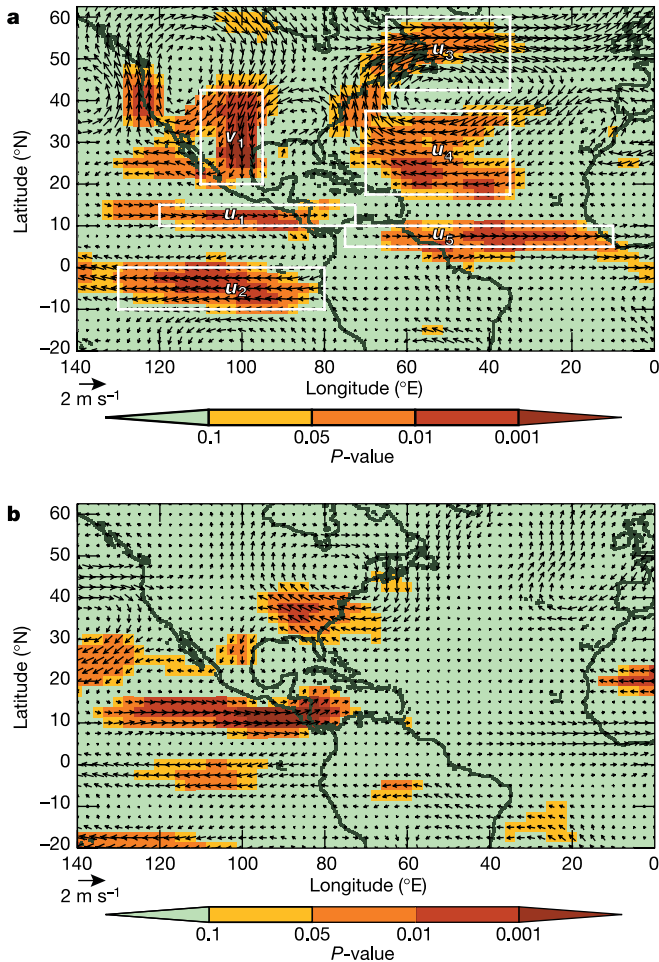


Figure 1 Tropospheric height-averaged wind anomalies linked significantly to above-median seasonal US landfalling hurricane activity 1950–2003. The panels show wind data for July (a) and August–October (b). Plotted is the difference in vector wind anomalies (averaged between 925 mbar and 400 mbar height) between those subset years when the US ACE index is in its upper and lower quartiles. The significance of the difference in wind magnitude between these subset years is shown by the colour bar. Seasonal predictability of the US ACE index is assessed using a July height-averaged wind index using the six regions marked by white boxes in a.

Table 1 Predictive skill for the seasonal US ACE index				
Predictor	Period	Hindcast skill and significance		
		MSSS (%)	r_{rank}	P -value
July 925–400 mbar wind index (all regions)	1950–2003	38	0.67	<0.001
	1950–76	40	0.65	0.001
	1977–2003	36	0.70	<0.001
July 925–400 mbar wind index ($u_1 - u_2 - v_1$)	1950–2003	32	0.68	<0.001
	1950–76	28	0.68	0.002
	1977–2003	38	0.75	<0.001
July 925–400 mbar wind index ($u_3 - u_4 + u_5$)	1950–2003	19	0.52	<0.001
	1950–76	22	0.55	0.004
	1977–2003	14	0.47	0.008
Observed Atlantic total ACE index	1950–2003	18	0.49	<0.001
	1950–76	13	0.38	0.03
	1977–2003	26	0.57	0.003
Observed August–October Niño 3 SST	1950–2003	0	0.29	0.05
	1950–76	0	0.26	0.12
	1977–2003	3	0.31	0.12

Table 2 Link between hindcast US ACE index and US hurricane losses				
Period	Economic loss		Insured loss	
	r_{rank}	P -value	r_{rank}	P -value
1950–2003	0.48	0.003	0.48	0.003
1950–76	0.43	0.06	0.58	0.03
1977–2003	0.49	0.008	0.37	0.03

(model 4) and of August–October Niño 3 SST (model 5) during the main hurricane season. The US ACE index hindcast skill is assessed using two skill measures: the correlation (r_{rank}) between the hindcast and observed values, and the mean square skill score (MSSS) defined as the percentage reduction in mean square error of the model hindcasts compared to hindcasts made with the 1950–2003 mean or climatology value. MSSS is the skill metric recommended by the World Meteorological Organisation for verification of deterministic seasonal forecasts²¹. P -values are computed from bootstrapped estimates of r_{rank} .

Table 1 shows the hindcast skill from the five models described above for different time periods. Using the July height-averaged wind index (all regions) model, the US ACE index (landfalling hurricane wind energy) is predictable from the 1 August start of the main Atlantic hurricane season with a correlation skill of 0.65–0.70 and a skill improvement over climatology of 36–40%. This skill is significant to $P < 0.001$ over the 1950–2003 period and to

$P = 0.001$ over each 1950–76 and 1977–2003 sub-period. The majority of this skill comes from 925–400 mbar winds over the east Pacific and North America. The strength, significance and stationarity of hindcast skill from the July wind model exceeds (by a factor of two in strength) that achievable from knowing the observed North Atlantic total ACE index at the hurricane season end on 30 November. The low hindcast skill shown by the August–October Niño 3 SST model confirms that ENSO-related wind anomalies are more important than ENSO SST anomalies for US hurricane seasonal predictability. A scatter plot of observed US ACE index versus hindcast US ACE index from the July wind index (all regions) model shows that it anticipates correctly 67% of upper and lower tercile actual values as being upper and lower tercile. Furthermore, the model anticipates correctly 94% of upper tercile actual values as above median and 83% of lower tercile actual values as below median.

The US ACE index hindcasts from the July wind index model offer sound potential for socio-economic benefit. Table 2 shows that the hindcasts are linked significantly ($P < 0.01$) and stably to US hurricane economic loss 1950–2003 (ref. 3) and to US hurricane insured loss 1950–2003 (ref. 17). For both economic and insured losses, $r_{\text{rank}} = 0.48$ and $P = 0.003$. These significant links to loss are evident also from Table 3, which compares the hindcast US ACE index values against economic and insured losses stratified by year and above/below median value. For economic loss, the hindcast model correctly anticipates the sign of 74% (40 out of 54) of the above-median and below-median loss years. For insured loss, the hindcast model correctly anticipates the sign of 70% (38 out of 54) of the above-median and below-median loss years. The two-tailed probability of obtaining the 2×2 contingency table of US ACE index hindcast and US economic loss by random chance is 0.001 based on Fisher's exact probability test; for hurricane insured loss the probability is 0.006.

A clear beneficiary of the above skill is the insurance industry. Buyers and sellers of reinsurance covers can improve returns over a period of years by up to 30% by using these forecasts²⁴. The July 925–400 mbar wind index model performed well in real-time operation in 2004, predicting a US ACE index in the upper quartile for this active and damaging hurricane season (see Atlantic forecast document dated 4th August 2004 at www.tropicalstormrisk.com). Insurers and others would have reduced their losses in 2004 by acting upon the forecast. □

Methods

Serial autocorrelation

We correct for serial autocorrelation to minimize the influence of time series trends and multi-year-to-decadal signal variability on the significance of the deduced hindcast skill^{18,19}. We do this by computing the effective number of degrees of freedom in the estimation of the cross-correlation by including autocorrelation coefficients in both input time series out to lags of $N/2$ yr, where N is the time series length.

Significances

The significances in Fig. 1 are computed for each grid cell by selecting two random composite sets of 14 values from the 54-yr 925–400 mbar height-averaged vector wind time series for that grid cell, and computing the wind magnitude difference between these random composites. The selection of values is made with replacement thereby allowing the same value to be picked more than once in a given set. This process is repeated 10,000 times for each grid cell. The random composite wind magnitude differences are displayed in histogram form to give the percentage of the random sets with a wind magnitude difference greater than the actual observed composite difference.

The significances in Tables 1 and 2 are computed by randomly shuffling the 54-yr hindcast/observed US ACE index time series and selecting with replacement a number of hindcast and observed US ACE index pairs corresponding to the number of degrees of freedom after correction for serial correlation in the unshuffled time series. This process is repeated 10,000 times. The correlation from each random set is calculated and the results are displayed in histogram form to give the percentage of the random sets with an r_{rank} greater than the original hindcast r_{rank} .

Predictor selection

The following predictor selection rule is used to select regions and wind components for the July height-averaged wind index. The July area-averaged wind anomaly component

Table 3 Hindcast US ACE index and annual US hurricane damage

Economic losses				Insured losses			
Year	Hindcast	Loss	US\$ million	Year	Hindcast	Loss	US\$ million
1992	—	+	44,014	1992	—	+	29,597
1954	+	+	23,302	1954	+	+	18,259
1955	+	+	17,548	1965	+	+	13,922
1965	+	+	16,888	1989	—	+	6,845
1960	+	+	16,236	1964	+	+	5,885
1969	+	+	14,584	1960	+	+	5,707
1972	—	+	14,258	1970	+	+	5,522
1989	—	+	13,705	1979	+	+	5,160
1979	+	+	11,489	1983	—	+	4,729
1961	+	+	9,536	1985	+	+	4,298
1964	+	+	9,377	1961	+	+	4,202
1985	+	+	8,834	1995	+	+	3,710
1999	—	+	6,346	1950	+	+	3,701
2001	+	+	5,579	1969	+	+	3,568
1983	—	+	5,395	1955	+	+	2,946
1995	+	+	4,957	2001	+	+	2,667
1996	+	+	4,635	1996	+	+	2,514
1970	+	+	4,439	1999	—	+	2,430
1998	+	+	4,414	1998	+	+	2,044
1950	+	+	3,732	2003	+	+	1,775
2003	+	+	3,580	1957	—	+	1,422
1957	—	+	3,251	1959	+	+	1,214
1967	+	+	2,726	1972	—	+	1,157
1975	+	+	2,336	1991	—	+	1,117
1991	—	+	2,279	1967	+	+	1,073
1971	+	+	1,612	1975	+	+	946
1994	+	+	1,367	2002	—	+	648
2002	—	—	1,244	1980	—	—	343
1980	—	—	1,151	1956	+	—	332
1974	—	—	953	1966	—	—	255
1959	+	—	594	1984	+	—	162
1956	+	—	466	1976	—	—	155
1968	—	—	425	1971	+	—	147
1976	—	—	408	1974	—	—	143
1958	—	—	296	1968	—	—	117
1951	+	—	242	1953	+	—	113
1966	—	—	219	1986	—	—	84
1963	+	—	197	1952	—	—	67
1984	+	—	173	1993	—	—	57
1973	—	—	126	1997	—	—	50
1997	—	—	123	1988	+	—	23
1988	+	—	116	1977	—	—	14
1981	—	—	102	1963	+	—	5
1978	—	—	100	1987	—	—	1
1990	—	—	99	1951	+	—	0
1993	—	—	85	1994	+	—	0
1952	—	—	84	1981	—	—	0
1962	—	—	56	1990	—	—	0
1977	—	—	44	1973	—	—	0
1986	—	—	39	1978	—	—	0
1953	+	—	37	2000	—	—	0
1982	—	—	36	1962	—	—	0
2000	—	—	30	1982	—	—	0
1987	—	—	18	1958	—	—	0

The comparison is made for the time period 1950–2003. Hindcasts are from the July height-averaged wind index model. Hurricane damage is shown separately in terms of economic and insured losses. The yearly value of each parameter is colour-coded on the basis of whether it is above-median (red, +) or below-median (blue, —). Rows are stratified vertically by loss and referenced by year. Losses are in millions of US dollars at 2003 prices and exposures.

must be linked significantly (correlation P -value <0.1 after correction for serial autocorrelation) to the US ACE index over each sub-period 1950–76 and 1977–2003, and for each year 1950 to 2003 after data excluding a 5-yr block centred on the year in question are excluded. This rule simulates the predictor selection process in an actual forecast situation.

Linear regression modelling

The US ACE index has a positively skewed (generalized Pareto) distribution. To satisfy the assumptions for using ordinary least squares regression, we transform this distribution to a normal distribution using the $\log(1 + \text{US ACE index})$ transform. We test for normality using the Kolmogorov–Smirnov test. The linear regression modelling is performed on these transformed data to produce hindcasts, which are then transformed back before the hindcast skill is computed. This procedure ensures that the observations are drawn from a normal distribution, and that the hindcast errors are normally distributed with a mean of zero (both requirements of linear regression modelling^{22,23}). A one-way analysis of variance (F -test) shows that the variance of the transformed observations and the variance of the hindcast errors are both constant in time (a further assumption of linear regression modelling^{22,23}).

The regression modelling is performed with a single predictor variable (the July wind index) rather than as a multiple regression with two or more predictor variables. Multiple regression is found always to give lower hindcast skill, the skill reduction increasing as the number of parameters increases. The single predictor (wind index) approach offers greater skill through better strengthening of the predictive signals and better removal of noise.

Cross-validated hindcasts

Cross-validated hindcasts are made with block elimination^{20,21}. The US ACE index of each year is hindcast by training the linear regression model on all data excluding a 5 yr block centred on the year of interest. The block is tapered at the time series ends. Block elimination is used to minimize potential skill inflation that might arise from multi-annual persistence. Cross-validation provides the best available estimate of forecast skill for the 54-yr sample.

Spearman rank correlation

The Spearman rank correlation coefficient is used as a robust and resistant alternative to the Pearson product–moment correlation coefficient²². The rank correlation is robust to deviations from linearity in a relationship, and is resistant to the influence of outliers.

Received 7 April 2004; accepted 31 January 2005; doi:10.1038/nature03454.

1. Benfield Industry Analysis and Research Team, *Catastrophe Losses* (Benfield, London, 2004).
2. Diaz, H. F. & Pulwarty, R. S. *Hurricanes: Climate and Socioeconomic Impacts* (Springer, Berlin, 1997).
3. Pielke, R. A. Jr & Landsea, C. W. Normalised hurricane damage in the United States: 1925–1995. *Weath. Forecast.* **13**, 621–631 (1998).
4. Lehmiller, G. S., Kimberlain, T. B. & Elsner, J. B. Seasonal prediction models for North Atlantic basin hurricane location. *Mon. Weath. Rev.* **125**, 1780–1791 (1997).
5. Gray, W. M., Landsea, C. W., Mielke, P. W. Jr & Berry, K. J. Predicting Atlantic basin seasonal tropical cyclone activity by 1 August. *Weath. Forecast.* **8**, 73–86 (1993).
6. Gray, W. M. Atlantic seasonal hurricane frequency. Part II: Forecasting its variability. *Mon. Weath. Rev.* **112**, 1669–1683 (1984).
7. Klotzbach, P. J. & Gray, W. M. Forecasting September Atlantic basin tropical cyclone activity. *Weath. Forecast.* **18**, 1109–1128 (2003).
8. Bove, M. C., Elsner, J. B., Landsea, C. W., Niu, X. & O'Brien, J. J. Effect of El Niño on U.S. landfalling hurricanes revisited. *Bull. Am. Meteorol. Soc.* **79**, 2477–2482 (1998).
9. Pielke, R. A. Jr & Pielke, R. A. Sr. La Niña, El Niño and Atlantic hurricane damages in the United States. *Bull. Am. Meteorol. Soc.* **80**, 2027–2034 (1999).
10. Saunders, M. A., Chandler, R. E., Merchant, C. J. & Roberts, F. P. Atlantic and NW Pacific typhoons: ENSO spatial impacts on occurrence and landfall. *Geophys. Res. Lett.* **27**, 1147–1150 (2000).
11. Lyons, S. W. U. S. tropical cyclone landfall variability 1950–2002. *Weath. Forecast.* **19**, 473–480 (2004).
12. Waple, A. M. et al. Climate assessment for 2001. *Bull. Am. Meteorol. Soc.* **83**, S1–S62 (2001).
13. Neumann, C. J., Jarvinen, B. R., McAdie, C. J. & Hammer, G. R. *Tropical Cyclones of the North Atlantic Ocean 1871–1998* (Historical Climatology Series 6–2, National Oceanic and Atmospheric Administration, Asheville, 1999).
14. Kalnay, E. et al. The NCEP/NCAR 40-year reanalysis. *Bull. Am. Meteorol. Soc.* **77**, 437–471 (1996).
15. Dong, K. & Neumann, C. J. The relationship between tropical cyclone motion and environmental geostrophic flows. *Mon. Weath. Rev.* **114**, 115–122 (1986).
16. Franklin, J. L., Feuer, S. E., Kaplan, J. & Aberson, S. Tropical cyclone motion and surrounding flow relationships: searching for beta gyres in Omega dropwindsonde datasets. *Mon. Weath. Rev.* **124**, 64–84 (1996).
17. Collins, D. J. & Lowe, S. P. A macro validation dataset for U.S. hurricane models. 217–252 (Casualty Actuarial Society, Winter Forum, 2001); available from CAS at (<http://www.casact.org/pubs>).
18. Davis, R. E. Predictability of sea surface temperatures and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.* **6**, 249–266 (1976).
19. Chen, W. Y. Fluctuations in northern hemisphere 700 mb height field associated with the Southern Oscillation. *Mon. Weath. Rev.* **110**, 808–823 (1982).
20. Elsner, J. B. & Schertmann, C. P. Assessing forecast skill through cross-validation. *Weath. Forecast.* **9**, 619–624 (1994).
21. Standardised verification system (SVS) for long-range forecasts (LRF). *New Attachment II-9 to the Manual on the GDPS (WMO-No. 485)*, Vol. 1 (WMO, Geneva, 2002).
22. Wilks, D. S. *Statistical Methods in the Atmospheric Sciences* (Academic, San Diego, 1995).
23. Freund, R. J. & Wilson, W. J. *Regression Analysis: Statistical Modeling of a Response Variable* (Academic, San Diego, 1998).
24. Hilti, N., Saunders, M. A. & Lloyd-Hughes, B. Forecasting stronger profits. *Glob. Reinsurance* 6–7 (July/August 2004).

Acknowledgements We thank B. Lloyd-Hughes for help with Table 3 and other aspects of the work. J. B. Elsner, C. W. Landsea, F. Vitart, J. Heming and I. M. Mason are thanked for comments on the manuscript. This work is supported by the TSR (Tropical Storm Risk) venture sponsored by Benfield (an independent reinsurance intermediary), Royal & SunAlliance (an insurance group), and Crawford & Company (a claims management solutions company). We acknowledge NOAA-CIRES, Climate Diagnostics Center, Boulder, Colorado, for the NCEP/NCAR Global Reanalysis Project data, and NOAA's Hurricane Research Division for the HURDAT North Atlantic hurricane database.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.A.S. (mas@mssl.ucl.ac.uk).

Active out-of-sequence thrust faulting in the central Nepalese Himalaya

Cameron Wobus¹, Arjun Heimsath², Kelin Whipple¹ & Kip Hodges¹

¹Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

²Department of Earth Sciences, Dartmouth College, Hanover, New Hampshire 03755, USA

Recent convergence between India and Eurasia is commonly assumed to be accommodated mainly along a single fault—the Main Himalayan Thrust (MHT)—which reaches the surface in the Siwalik Hills of southern Nepal^{1–3}. Although this model is consistent with geodetic^{4,5}, geomorphic⁶ and microseismic data⁷, an alternative model incorporating slip on more northerly surface faults has been proposed to be consistent with these data as well^{8–10}. Here we present *in situ* cosmogenic ¹⁰Be data indicating a fourfold increase in millennial timescale erosion rates occurring over a distance of less than 2 km in central Nepal, delineating for the first time an active thrust fault nearly 100 km north of the surface expression of the MHT. These data challenge the view that rock uplift gradients in central Nepal reflect only passive transport over a ramp in the MHT. Instead, when combined with previously reported ⁴⁰Ar–³⁹Ar data⁹, our results indicate persistent exhumation above deep-seated, surface-breaking structures at the foot of the high Himalaya. These results suggest that strong dynamic interactions between climate, erosion and tectonics have maintained a locus of active deformation well to the north of the Himalayan deformation front.

The central Nepalese Himalaya is a textbook example of continent–continent collision, in which the underthrusting of India has been concentrated on several roughly east–west-trending fault zones within a belt about 100 km wide. The northernmost of these fault zones is the Main Central Thrust (MCT), which marks a transition from the high-grade metamorphic Greater Himalayan Sequence in the north to the lower-grade Lesser Himalayan Sequence in the south. Geochronologic data indicate that the MCT is also the oldest structure, with evidence for initial activity on this thrust fault by 23–20 Myr ago¹¹. More southerly structures—the Main Boundary Thrust (MBT) and the Main Frontal Thrust (MFT)—developed progressively in a north–south sequence, consistent with observations in foreland fold and thrust belts worldwide¹² (Fig. 1a). Most researchers working in the Nepal orogen assume that recent surface faulting has been concentrated at the trace of the MFT, which defines the southern limit of deformation in the Himalayan system. In this model, the MFT absorbs almost all slip on the MHT. However, this interpretation does not provide a

direct explanation for the striking contrast between high modern surface uplift rates in the high Himalayan ranges and the much lower rates in the Himalayan foothills⁴, which occurs nearly 100 km north of the MFT across a distinctive physiographic transition. It has been suggested that these changes in physiography and surface uplift rate are best explained by a gradual ramp in the MHT in the middle crust^{1,4–6} (Fig. 1b).

While the ramp hypothesis is consistent with most geological and geophysical data from the Nepalese Himalaya, both the sharpness of the physiographic transition in central Nepal and the relatively abrupt change in surface uplift rate across it are difficult to reconcile with the broader transitions that might be expected as manifestations of a midcrustal ramp¹. In the Burhi Gandaki valley, for example, mean elevation and relief (as measured along a 20-km-wide swath profile) each increase more than twofold over a distance of less than 8 km (Fig. 2). The lower boundary of this physiographic transition can be precisely delineated on the basis of changes in valley morphology, hillslope gradients, channel gradients and the extent of thick alluvial fill deposits⁹. Importantly, this lower boundary occurs between 20 and 30 km south of the surface trace of the MCT, indicating that an unmapped thrust fault might be accommodating gradients in rock uplift in this valley. Farther west in the Marsyandi valley, changes in landscape morphology occur more gradually, which is consistent with a more broadly distributed strain field. In the Marsyandi drainage, the upper boundary of the physiographic transition is nearly coincident with the mapped trace of the MCT, indicating that strands of the MCT itself might be important in accommodating modern rock uplift gradients in

this valley¹⁰. These along-strike differences indicate spatial variations in how recent deformation is accommodated along the Himalayan front. Nonetheless, the physiographic data from central Nepal are all broadly consistent with independent evidence for recent deformation in the region including young, brittle shear zones near the MCT in the Marsyandi valley¹⁰ and sharp discontinuities in the patterns of Late Miocene–Quaternary ⁴⁰Ar–³⁹Ar and fission-track mineral cooling ages throughout central Nepal^{9,13,14}.

Unfortunately, although the physiographic transition is sharp and well defined in the Burhi Gandaki river, the poor quality of bedrock exposure makes it difficult to construct detailed structural maps to determine unequivocally whether young, surface-breaking faults are present. Here we report the results of a different approach to the problem based on deducing differences in erosional patterns from cosmogenic radionuclide data in detrital sediments. We measured concentrations of ¹⁰Be produced *in situ* in modern sediment from eight small tributaries to the Burhi Gandaki as a proxy for millennial timescale erosion rates in each catchment. The concentration of ¹⁰Be in quartz, interpreted with nuclide production rates scaled for altitude, latitude and local topography, enables erosion rates to be quantified at the outcrop scale¹⁵. At the basin scale, ¹⁰Be concentrations in sediment have been shown to represent reliable basin-average erosion rates in a variety of climatic and tectonic settings^{15–19}. We use the spatial pattern of erosion rates from the Burhi Gandaki tributaries to delineate discontinuities in rock uplift rates across the range front. Field observations indicate that hillslope angles might approach threshold values near the physiographic transition, so landscape response should be rapid: as bedrock rivers adjust their incision rates in response to spatial variations in rock uplift, we expect the hillslopes to match this incision rate by landsliding to maintain their critical condition^{20,21}.

The calculation of a basin-average erosion rate from ¹⁰Be concentrations in sediment from steep, landslide-dominated catchments requires an assumption that the sediment collected at the basin outlet is well mixed, and therefore that pulses of cosmogenically ‘underexposed’ landslide-derived material are integrated into the bulk sediment sample. Larger basins will more effectively integrate these stochastic sediment pulses downstream, indicating that basin scale might be an important factor in controlling the fidelity of the cosmogenic signal^{22–24}. Basins sampled for this study have drainage areas ranging from about 3 km² to about 22 km² (Table 1), a range in which preliminary numerical modelling suggests a high probability that basin-average erosion rates will be closely predicted—or only slightly underestimated—from detrital cosmogenic radionuclide concentrations²⁴. Furthermore, the drainage pattern in the Burhi Gandaki is trellised, with tributaries draining narrow (about 2–5 km wide) catchments subparallel to the structural grain of the orogen (Fig. 2a). Sediment from each tributary therefore records an estimate of the basin-average erosion rate from a narrowly constrained tectonostratigraphic position.

The ¹⁰Be data reveal a sharp discontinuity in erosion rates centred about 23 km south of the MCT, within the zone of the physiographic transition as defined by independent methods⁹. To the south of this discontinuity, erosion rates are uniform at about 0.2 mm yr^{–1}. To the north, erosion rates abruptly jump to about 0.8 mm yr^{–1} and then gradually decline to about 0.2 mm yr^{–1} over a distance of 10 km (Fig. 3a). ¹⁰Be concentrations (Table 1) indicate minimum exposure ages ranging from about 1 kyr to 3 kyr, indicating sharp spatial gradients in basin-average erosion rates over late Holocene timescales. This break in erosion rates is not correlated with any mappable break in lithology: rocks to the north and south of this transition each comprise phyllites and schists of the Lesser Himalayan sequence. Furthermore, there is no correlation between the cosmogenically determined erosion rates and basin size, and the fourfold increase in erosion rates is larger than any bias that might be predicted by preliminary modelling accounting for under-sampling of landslide-derived material in small catchments²⁴. The

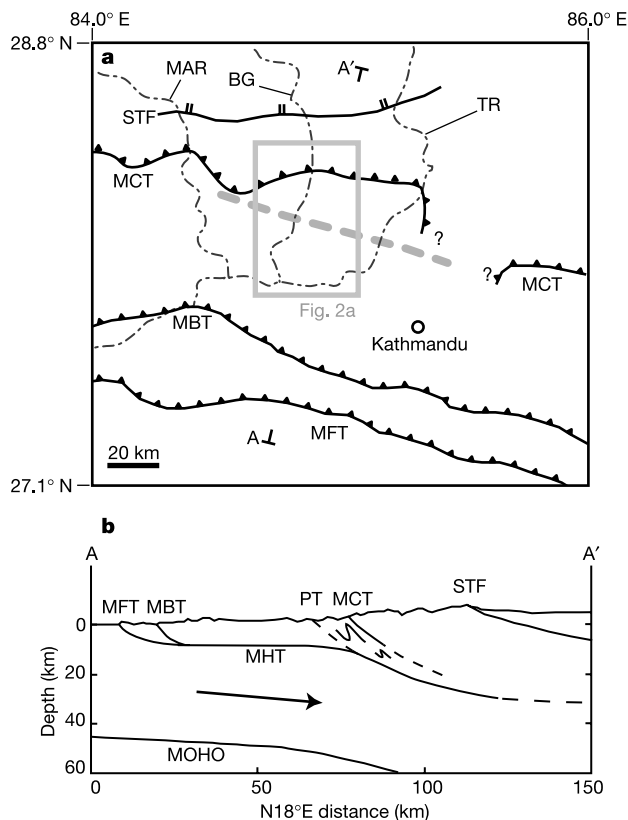


Figure 1 Geological setting. **a**, Regional geological map showing major tectonic structures and river systems. Dash-dotted lines: MAR, Marsyandi river; BG, Burhi Gandaki river; TR, Trisuli river. STF, South Tibetan fault. Dashed grey line shows the lower boundary of the physiographic transition where it is well defined (see the text for explanation). The grey box indicates the location of Fig. 2a. A–A' indicates the location of schematic cross section in **b**. **b**, Schematic cross section across central Nepal, showing the ramp in the MHT and the inferred projection of a fault at the physiographic transition (PT).

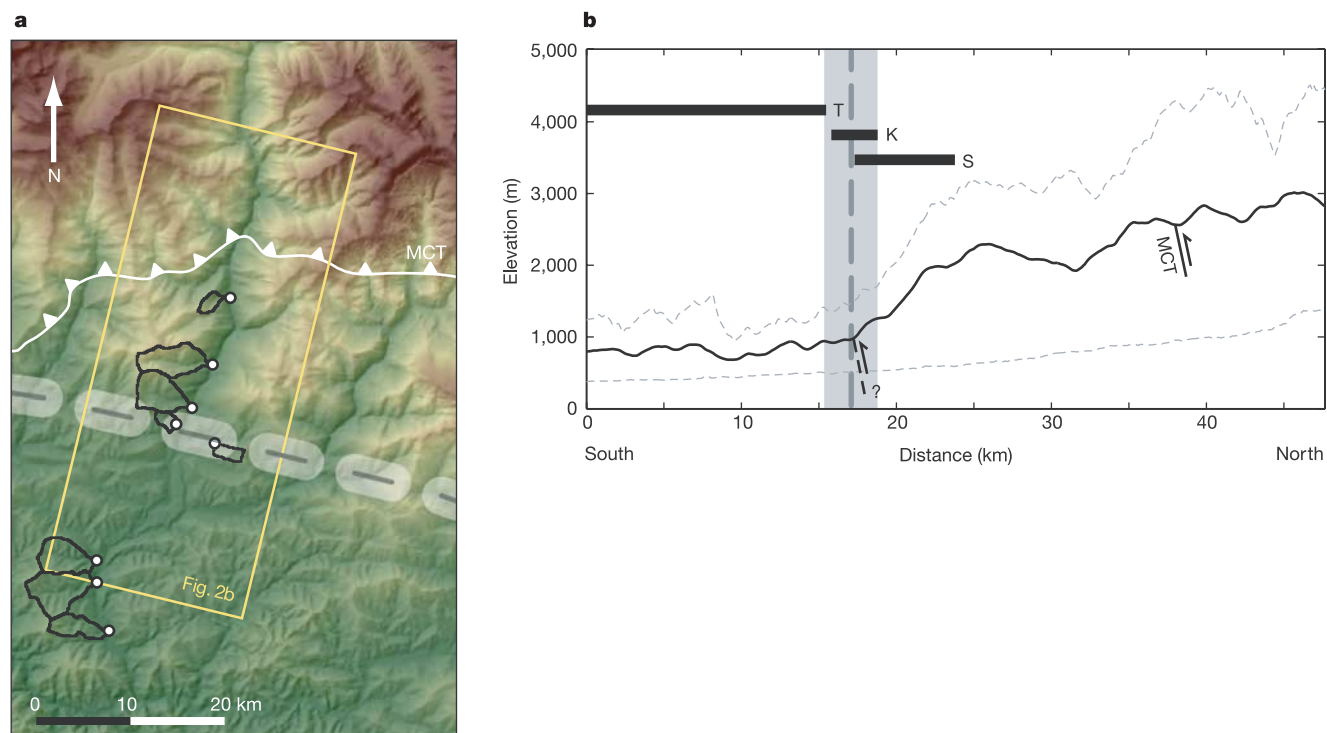


Figure 2 Sampling locations and study area physiography. **a**, Sample collection points (white dots) and drainage basins (black outlines). The dashed grey line shows approximate trace of the physiographic transition. The yellow rectangle delineates the boundaries of the swath profile shown in **b**. **b**, Mean (black line), minimum and maximum elevations

spatial trend in ^{10}Be erosion rates therefore corroborates our interpretations of tectonics based solely on landscape morphology, and allows us to constrain more narrowly the locus of active thrust faulting in the Burhi Gandaki valley.

The discontinuity in short-term erosion rates is also co-located with a prominent break in long-term cooling rates from thermochronology (Fig. 3b): to the north of the physiographic transition, muscovite ^{40}Ar – ^{39}Ar cooling ages are young (Cenozoic), whereas they are substantially older to the south (Palaeozoic–Proterozoic)⁹. Calculation of long-term exhumation rates from these data would require a detailed thermal model accounting for lateral advection of rock and temporal variations in the subsurface thermal structure. Although this calculation is beyond the scope of this study, the abrupt discontinuity in cooling ages is a robust finding that corroborates our interpretation of a surface-breaking thrust fault at the physiographic transition. Furthermore, the thermochronologic data provide the additional constraint that active thrust faulting has persisted at least long enough to create a substantial discontinuity in the total depth of exhumation from north to south.

The presence of a tectonically significant, thrust-sense fault zone at the physiographic transition, as implied by the spatial

(dashed grey lines) along a 20-km-wide swath profile oriented orthogonal to the strike of the range. The vertical grey line marks the base of the physiographic transition (see text). Black bars show the extent of alluvial fill terraces (T), knickpoints on Burhi Gandaki tributaries (K) and the zone of increasing steepness on the Burhi Gandaki trunk stream (S).

coincidence of breaks in short-term (cosmogenic) and long-term (^{40}Ar – ^{39}Ar) erosion rates, is consistent with field observations of brittle deformational fabrics parallel to the northward-dipping foliations in the Lesser Himalayan Sequence⁹. Physiographic data from along strike suggest that this fault zone extends eastwards to the Trisuli valley, maintaining its position substantially south of the MCT. To the west of the Burhi Gandaki valley, the orientation of the physiographic transition and the more diffuse gradients in landscape morphology suggest that this fault zone becomes more broadly distributed in the Marsyandi valley, and might correspond to recent activity within the MCT zone^{10,14}.

We speculate that the origin of this fault zone might be intimately tied to the presence of strong precipitation gradients across the central Nepalese Himalaya. Focused monsoonal precipitation is well documented on the southern flank of the Nepalese Himalaya^{14,25} and has been posited to drive localized tectonic uplift by removing mass from the top of an extruding ductile channel^{26,27}. Although the energy driver for this channel extrusion is gravitational potential energy from the Tibetan plateau, erosion must be important in determining where the energy is dissipated. Our data suggest there might be a dynamic feedback between climate and tectonics in the

Table 1 Basin characteristics and cosmogenic erosion rate data

Sample	Distance from MCT (km)*	Drainage area (km ²)	Mean slope (deg)†	Elevation range (m)‡	Mass of quartz (g)	[^{10}Be] (10^3 atoms g ⁻¹)	Erosion rate (mm yr ⁻¹)
01WBS5	7.0–10.0 (7.5)	3.4	28.4	797–2,372	58.64	42.1 ± 2.3	0.19 ± 0.02
01WBS6	13.5–17.5 (15.0)	18.4	30.8	604–3,158	79.32	27.8 ± 1.6	0.37 ± 0.04
01WBS7	17.0–21.5 (19.5)	17.5	24.6	533–2,455	69.25	13.9 ± 1.7	0.48 ± 0.08
03WBS1	21.0–23.0 (22.0)	3.2	21.4	643–1,325	150.90	6.0 ± 0.5	0.77 ± 0.10
03WBS2	22.0–24.0 (23.0)	3.9	17.6	723–1,475	150.24	27.9 ± 0.9	0.19 ± 0.01
01WBS3	37.0–41.5 (38.5)	16.7	22.0	413–1,412	67.14	21.9 ± 1.9	0.19 ± 0.03
01WBS2	40.0–46.0 (41.0)	22.4	21.9	370–1,574	70.03	23.4 ± 1.8	0.19 ± 0.02
01WBS1	44.5–47.0 (45.5)	10.5	21.4	348–1,670	84.88	25.4 ± 1.7	0.18 ± 0.02

Errors are ±1σ.

*Distance to the northern and southern edges of the basin, rounded to the nearest 0.5 km and projected onto a line oriented N18°E. Numbers in parenthesis are distances to the basin outlet.

†Slopes calculated from a 3 × 3 moving window over 90-m resolution digital elevation model (DEM).

‡Production rates calculated pixel by pixel by using 90-m DEM.

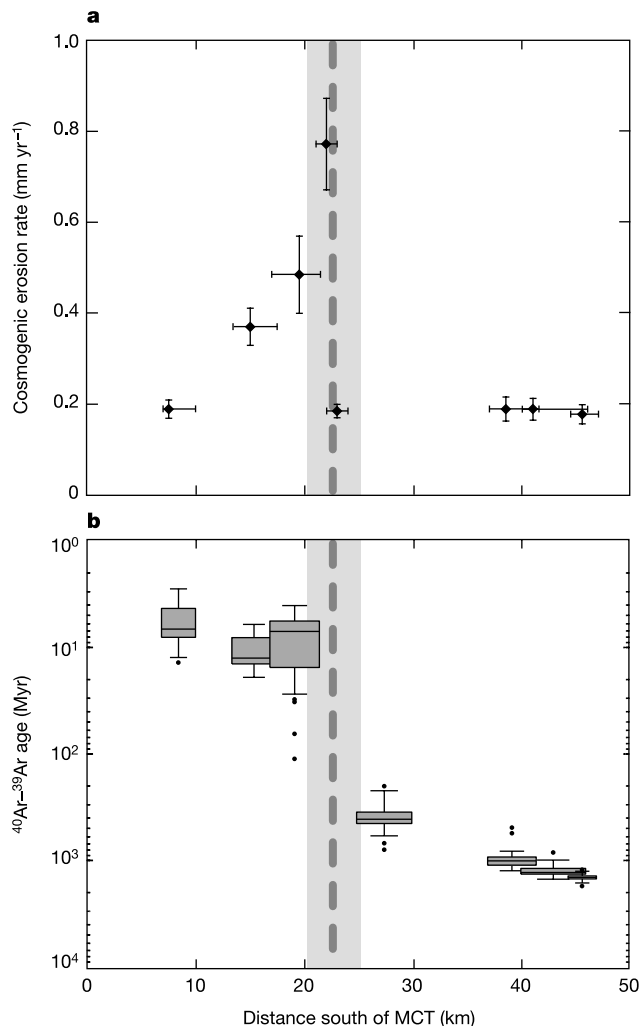


Figure 3 Erosion rate and cooling-age data. **a**, Plot of ¹⁰Be erosion rate (dots) against distance from the MCT, projected onto a N18°E line. Error bars on the x axis represent the projected distance to basin limits; error bars on the y axis represent 1σ uncertainty in analytical results. **b**, Plot of ⁴⁰Ar-³⁹Ar cooling ages (logarithmic scale) against distance from the MCT. Shaded boxes, vertical whiskers, and horizontal lines within boxes represent the interquartile range, limits of analysis results, and median values, respectively. Black dots represent outliers (more than 1.5 times the interquartile range beyond box limits). Widths of boxes represent widths of individual basins. Complete data can be found in ref. 9. Vertical shading and dashed lines show physiographic transition.

Himalayan orogen, so much so that the locus of deep exhumation has been maintained nearly 100 km northwards of the Himalayan thrust front. This focused exhumation sustains the marked topographic front of the high Himalaya, and increases the efficiency of energy dissipation from the Himalayan system. □

Received 14 October 2004; accepted 23 February 2005; doi:10.1038/nature03499.

- Cattin, R. & Avouac, J. P. Modeling mountain building and the seismic cycle in the Himalaya of Nepal. *J. Geophys. Res.* **105**, 13389–13407 (2000).
- Lave, J. & Avouac, J. P. Active folding of fluvial terraces across the Siwaliks Hills, Himalayas of central Nepal. *J. Geophys. Res.* **105**, 5735–5770 (2000).
- Pandey, M. R., Tandukar, R. P., Avouac, J. P., Lave, J. & Massot, J. P. Interseismic strain accumulation on the Himalayan crustal ramp (Nepal). *Geophys. Res. Lett.* **22**, 751–754 (1995).
- Jackson, M. & Bilham, R. Constraints on Himalayan deformation inferred from vertical velocity fields in Nepal and Tibet. *J. Geophys. Res.* **99**, 13897–13912 (1994).
- Bilham, R. *et al.* GPS measurements of present-day convergence across the Nepal Himalaya. *Nature* **386**, 61–64 (1997).
- Lave, J. & Avouac, J. P. Fluvial incision and tectonic uplift across the Himalayas of central Nepal. *J. Geophys. Res.* **106**, 26561–26591 (2001).
- Pandey, M. R. *et al.* Seismotectonics of the Nepal Himalaya from a local seismic network. *J. Asian Earth Sci.* **17**, 703–712 (1999).

- Catlos, E. J. *et al.* Geochronologic and thermobarometric constraints on the evolution of the Main Central Thrust, central Nepal Himalaya. *J. Geophys. Res.* **106**, 16177–16204 (2001).
- Wobus, C. W., Hodges, K. V. & Whipple, K. X. Has focused denudation sustained active thrusting at the Himalayan topographic front? *Geology* **31**, 861–864 (2003).
- Hodges, K., Wobus, C., Ruhl, K., Schildgen, T. & Whipple, K. Quaternary deformation, river steepening and heavy precipitation at the front of the Higher Himalayan ranges. *Earth Planet. Sci. Lett.* **220**, 379–389 (2004).
- Hodges, K. V. Tectonics of the Himalaya and southern Tibet from two perspectives. *GSA Bull.* **112**, 324–350 (2000).
- DeCelles, P. G. *et al.* Stratigraphy, structure, and tectonic evolution of the Himalayan fold-thrust belt in western Nepal. *Tectonics* **20**, 487–509 (2001).
- Copeland, P. *et al.* An early Pliocene thermal disturbance of the Main Central Thrust, central Nepal; implications for Himalayan tectonics. *J. Geophys. Res.* **96**, 8475–8500 (1991).
- Burbank, D. W. *et al.* Decoupling of erosion and climate in the Himalaya. *Nature* **426**, 652–655 (2003).
- Bierman, P. R. & Nichols, K. K. Rock to sediment-slope to sea with ¹⁰Be-rates of landscape change. *Annu. Rev. Earth Planet. Sci.* **32**, 215–255 (2004).
- Brown, E. T., Stallard, R. F., Larsen, M. C., Raisbeck, G. M. & Yiou, F. Denudation rates determined from the accumulation of in-situ produced ¹⁰Be in the Luquillo Experimental Forest, Puerto Rico. *Earth Planet. Sci. Lett.* **129**, 193–202 (1995).
- Riebe, C. S., Kirchner, J. W., Granger, D. E. & Finkel, R. C. Erosional equilibrium and disequilibrium in the Sierra Nevada, inferred from cosmogenic ²⁶Al and ¹⁰Be in alluvial sediment. *Geology* **28**, 803–806 (2000).
- Schaller, M., von Blanckenburg, F., Hovius, N. & Kubik, P. W. Large-scale erosion rates from in-situ produced cosmogenic nuclides in European river sediments. *Earth Planet. Sci. Lett.* **188**, 441–458 (2001).
- Vance, D., Bickle, M., Ivy-Ochs, S. & Kubik, P. W. Erosion and exhumation in the Himalaya from cosmogenic isotope inventories in river sediments. *Earth Planet. Sci. Lett.* **206**, 273–288 (2003).
- Burbank, D. W. *et al.* Bedrock incision, rock uplift and threshold hillslopes in the northwestern Himalayas. *Nature* **379**, 505–510 (1996).
- Roering, J. J., Kirchner, J. W. & Dietrich, W. E. Hillslope evolution by nonlinear, slope-dependent transport: steady state morphology and equilibrium adjustment timescales. *J. Geophys. Res.* **106**, 16499–16513 (2001).
- Bierman, P. & Steig, E. Estimating rates of denudation using cosmogenic isotope abundances in sediment. *Earth Surf. Processes Landforms* **21**, 125–139 (1996).
- Granger, D. E., Kirchner, J. W. & Finkel, R. Spatially averaged long-term erosion rates measured from in situ-produced cosmogenic nuclides in alluvial sediment. *J. Geol.* **104**, 249–257 (1996).
- Niemi, N. A., Oskin, M. E. & Burbank, D. A numerical simulation of the effects of mass-wasting on cosmogenically determined erosion rates. *Eos* **85** (Fall Meet. Suppl.), Abstract H51C-1157 (2004).
- Putkonen, J. Continuous snow and rain data at 500 to 4400 m altitude near Annapurna, Nepal, 1999–2001. *Arct. Antarct. Alpine Res.* **36**, 244–248 (2004).
- Beaumont, C., Jamieson, R. A., Nguyen, M. H. & Lee, B. Himalayan tectonics explained by extrusion of a low-viscosity crustal channel coupled to focused surface denudation. *Nature* **414**, 738–742 (2001).
- Hodges, K. V., Hurtado, J. M. & Whipple, K. X. Southward extrusion of Tibetan crust and its effect on Himalayan tectonics. *Tectonics* **20**, 799–809 (2001).

Acknowledgements We thank D. Burbank and P. Bierman for constructive reviews, which greatly improved the quality of the original manuscript, R. Finkel at LLNL for accommodating our samples at short notice, and B. Crosby, K. Ruhl, T. Schildgen, N. Wobus and Himalayan Experience for field assistance. The work was funded by NSF and NSF Continental Dynamics.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to C.W. (cwobus@mit.edu).

Geobiology of a microbial endolithic community in the Yellowstone geothermal environment

Jeffrey J. Walker, John R. Spear & Norman R. Pace

Department of Molecular, Cellular and Developmental Biology and the Center for Astrobiology, University of Colorado, Boulder, Colorado 80309-0347, USA

The endolithic environment, the pore space of rocks, is a ubiquitous habitat for microorganisms on the Earth¹ and is an important target of the search for life elsewhere in the Solar System². Photosynthetic, endolithic microbial communities commonly inhabit the outer millimetres to centimetres of all rocks exposed to the Earth's surface. In the most extreme terrestrial climates, such as hot and cold deserts, endolithic microorganisms are often the main form of life^{3–5}. The endolithic microhabitat

gives protection from intense solar radiation and desiccation, and it provides mineral nutrients, rock moisture and growth surfaces^{4,5}. Here we describe the discovery and identification of the constituents of an extremely acidic (pH 1) endolithic microbial community inhabiting the pore space of rocks in the geothermal environment of Yellowstone National Park, USA. Subjected to silica mineralization, such endolithic communities constitute biomarkers that can become fossilized and potentially preserved in the geological record. Remnants of these communities could serve as biosignatures and provide important clues about ancient life associated with geothermal environments on the Earth or elsewhere in the Solar System.

Lush and unusual photosynthetic communities inhabit silica rocks in Yellowstone's Norris Geyser Basin. These rocks are primarily chalcedonic sinters and are warmed to ~35 °C by subsurface geothermal activity. The stark, weathered surfaces of these exposed rocks show no evidence of the rich life hidden beneath the surface (Fig. 1). Fractured rock samples show clear signs of photosynthetic endolithic communities, which inhabit a distinct green band from 1 to 15 mm thick and 2–10 mm beneath the surface exposed to light. Photosynthetic pigments, primarily chlorophyll of red algae (see below), impart a green colour. Although endolithic growth of red algae is known in other volcanic areas⁶, this is the first comprehensive molecular analysis of the microbial diversity and composition of these unique communities and their potential mineralization and fossilization.

Pore waters extracted from the rock had a pH of ~1 and contained high concentrations of sulphuric acid, metals and silicates (for example 1,343 p.p.m. SO₄, 8,893 p.p.m. Al, 983 p.p.m. Mn, 551 p.p.m. Cu, 1,038 p.p.m. As, 102 p.p.m. Fe and 166 p.p.m. SiO₂). The acidity is a product of the physical and biological oxidation of reduced sulphur species to sulphuric acid^{7,8}. Rocks that harboured endolithic communities contained 10–32% pore water by weight. The evaporation of pore water concentrates dissolved species, particularly silica, which precipitates and encrusts the microbial community as it grows. Similar mineralization processes are observed in living microbial communities directly associated with terrestrial^{2,9–11} and marine¹² thermal springs. The processes that fossilize microbial communities, how these fossils are preserved in the geological record, and their potential as biosignatures have been examined extensively^{2,9,13–16}. The community described here is a unique opportunity to observe an endolithic microbial community in the process of fossilization.

To understand the kinds of organism that constitute this previously unknown community, we used DNA sequence-based methods that rely on genomic DNA extracted directly from the environment instead of cultivation techniques that typically grow much less than 1% of the organisms present¹⁷. We isolated total community DNA from rock samples and used this as a template to amplify by polymerase chain reaction (PCR) small-subunit ribosomal RNA genes representative of organisms present in the community. We used universal PCR primers that target all three domains of life, the Bacteria, Archaea and Eukarya, as well as primers that target only the Bacteria. Individual PCR-amplified rRNA genes were cloned and segregated into libraries, collections of 96 randomly selected rRNA gene clones. In total, we screened and analysed 864 clones, of which we sequenced 342. Results (Fig. 2a) show that the community is relatively simple, composed mainly of organisms with rRNA gene sequences of ~40 species-level relatedness groups (at least 97% rRNA sequence identity), a conservative estimate of species-level diversity based on rRNA sequence variation¹⁸.

Two kinds of sequence dominated the clone libraries. About 26% of clones (225 of 864) were most closely related by phylogenetic analysis (98–99% sequence identity) to chloroplast rRNA genes of *Cyanidium caldarium*, an acidophilic eukaryotic red alga. This indicates that closely related *Cyanidium* species might be the

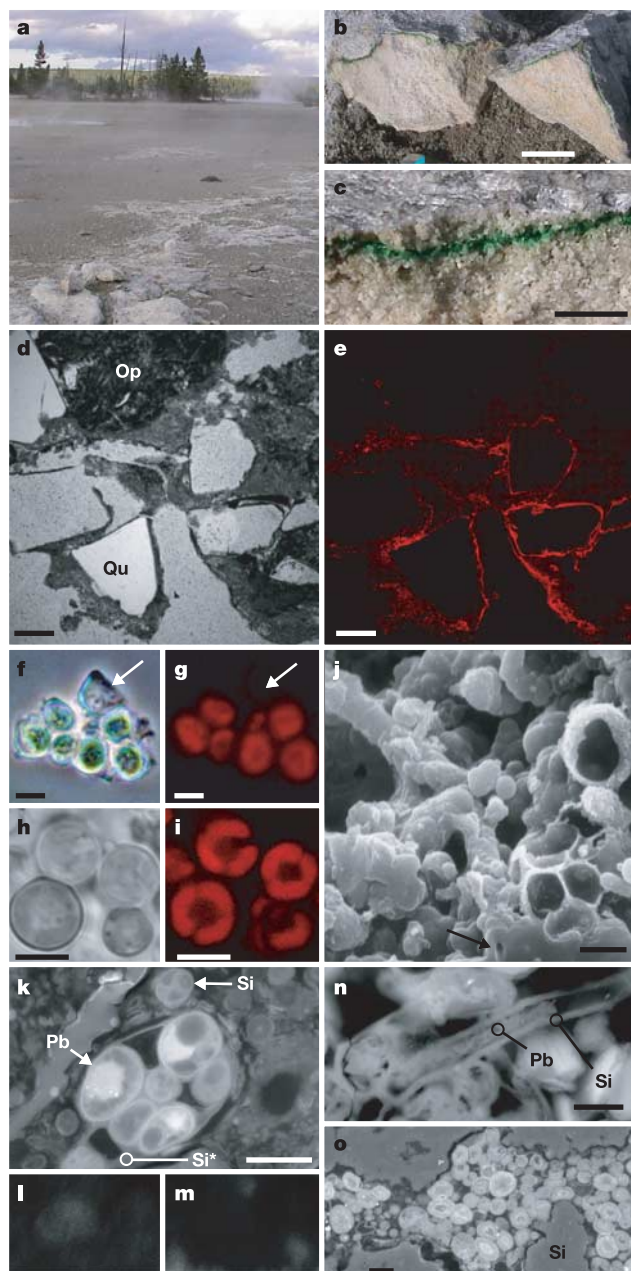


Figure 1 Endolithic microbial communities in highly acidic (pH1) chalcedonic sinters of Yellowstone's Norris Geyser Basin. **a**, Outcrops harbouring cryptoendoliths (foreground). **b**, **c**, Fractured sinter reveals distinct green layer of photosynthetic microbial life. Scale bars, 5 cm (**b**) and 1 cm (**c**). **d**, **e**, LSCM images of 30- μ m petrographic thin section of endolithic community. **d**, Transmitted light shows that host rock is chalcedonic sinter with quartz grains (Qu) cemented by microcrystalline opaline silica (Op). **e**, Fluorescence associated with cells containing chlorophyll imaged with LSCM (excitation at 488 nm, detection at 630–650 nm). Scale bars, 200 μ m. **f**, **g**, Micrographs of *Cyanidium* spp. encrusted in mineral casts. Transmitted light (**f**) shows casts with and without (arrows) fluorescent *Cyanidium* spp. cells (**g**). Scale bars, 5 μ m. **h**, **i**, LSCM micrographs of *Cyanidium* spp. cultivated in enrichments imaged with transmitted light (**h**) and fluorescence (**i**) as in **e**. Scale bars, 5 μ m. **j**, SEM image of broken spherical silica casts that form around individual *Cyanidium* spp. cells and a tubular cast (arrow) that forms around a network of cells presumed to be largely *Mycobacterium* spp. Scale bar, 5 μ m. **k–m**, BSE-SEM image (**k**) and EDS maps (**l**, **m**) of transverse section of presumed *Cyanidium* spp. cellular structures. Si*, EDS point spectra. Heavy metals stain biological material (Pb) (**l**). One cellular structure (Si) did not stain with Pb and had a strong Si signal (**m**). Scale bar, 5 μ m. **n**, BSE-SEM image of tubular structure. EDS point spectrum showed that the cast (Si) was primarily silicon and that the inner material (Pb) was stained with lead. Scale bar, 5 μ m. **o**, BSE-SEM image of cells in pore space surrounded by silica cement (Si). Scale bar, 5 μ m.

main primary producers of the community. *Cyanidium* and other undetected genera of the Cyanidiaceae family, *Cyanidioschyzon* and *Galdieria*, are the most acid-tolerant photosynthetic organisms known. Nuclear rRNA genes of *Cyanidium* spp. were detected in enrichment cultures of *Cyanidium* spp. but not in universal community libraries. Analysis of photosynthetic pigments extracted directly from the community showed only chlorophyll *a* and other associated photosynthetic pigments identical with those in *Cyanidium* species cultured from Norris Geyser Basin (see Methods).

The most abundant group of rRNA gene clones analysed, a remarkable 37% of the total (328 of 864), were previously unidentified *Mycobacterium* species, members of the bacterial division Actinobacteria (Fig. 2b). Enrichments were not successful in the cultivation of *Mycobacterium* spp., as assayed by rRNA probe hybridization. Famous as pathogens, mycobacteria cause diseases such as tuberculosis and leprosy in humans. Mycobacteria also are considered common constituents of environmental settings such as soil and freshwater habitats, but they have not previously been encountered as such a large component of an environmental

microbial community. Such organisms typically are of low abundance (much less than 1%) and detected in the environment only with specific and sensitive assays¹⁹. The role of mycobacteria in these natural settings is not well understood but is generally considered heterotrophic. Environmental mycobacteria can adapt to acidic environments. Studies of environmental soils and sediments with quantitative cultures show an enrichment of mycobacteria at low pH. The richest mycobacterial environment previously reported (~1% of the community) is a mildly acidic (pH 4) forest soil¹⁹. Mycobacteria have not been encountered previously in such highly acidic (pH 1) volcanic environments as those reported here.

Although the primary energy source for the community is probably photosynthesis, the flux of geochemical species through the habitat indicates that lithotrophy, the metabolism of reduced inorganic compounds for energy, might contribute to the overall energy budget of the community. Dissolved species that could potentially serve as fuels include metals, sulphur compounds and hydrogen, an important energy source for microbes abundant in Norris Geyser Basin²⁰. The most diverse group of rRNA gene sequences in the community represented ~14 distinct groups of the bacterial division Proteobacteria and comprised ~11% of the total clones. Most of these sequences (~9%) belong to the γ -Proteobacteria, a group that includes lithotrophic organisms that metabolize hydrogen and sulphide. In addition, rRNA gene sequences closely related (99% identity) to those of two known lithoautotrophs, *Sulfobacillus disulfidooxidans* and *Leptospirillum ferrooxidans*, each comprised ~1% of the community. Only ~5% of clones were representative of members of the domain Archaea, which are popularly associated with such 'extreme' geothermal environments; however, bacteria, and to some extent eukaryotes, are equally important in these settings¹⁷.

We used several imaging techniques to investigate the physical structure of the community. Light microscopy of petrographic thin sections showed the rock to be a chalcedonic sinter with large, weathered, quartz grains cemented by opaline silica (Fig. 1d). To determine the distribution of photosynthetic organisms *in situ*, we imaged the fluorescence associated with chlorophyll in thin sections. Fluorescence localized to the microcrystalline cement and was most intense along quartz grain margins (Fig. 1e). Individual cells were observed at higher magnifications, although the fluorescence of cells embedded in the mineral matrix was highly scattered. Images of *Cyanidium* cells in freshly disrupted samples showed that many apparently healthy cells were encrusted in a crystalline matrix, presumably silica (Fig. 1f, g). Spectral analysis of fluorescence in thin sections and that of individual *Cyanidium* spp. cells yielded identical results.

Field-emission scanning electron microscopy (SEM) revealed a complex fabric of mineralized filamentous and spherical casts (Fig. 1j) in the pigmented zone of colonization. Nearly identical mineralized silica biofabrics occur in living microbial mats (nominally *Cyanidium caldarium*) in acidic (pH 2) thermal springs in Hokkaido, Japan¹⁰. Similar fabrics are also observed in mineralized and recently fossilized microbial communities in other thermal spring environments^{2,9,11,12}. Recent findings indicate that fossilized subsurface biofabrics might be preserved in the geological record². We imaged sections of resin-embedded endolithic community with backscattered-electron-imaging scanning electron microscopy (BSE-SEM), which showed spherical and filamentous cellular structures embedded in the mineral matrix (Fig. 1k-m). Samples were treated with heavy metals (osmium, lead and uranium), which stain and give contrast to biological materials in BSE-SEM images. The spherical cellular structures had outer diameters of 3.5–6.6 μm (average $5.0 \pm 0.8 \mu\text{m}$, $n = 50$), similar to the size of *Cyanidium* cells we observed by light microscopy. We analysed the elemental composition of samples with energy-dispersive X-ray spectroscopy (EDS). The interiors of most cellular structures had strong signals

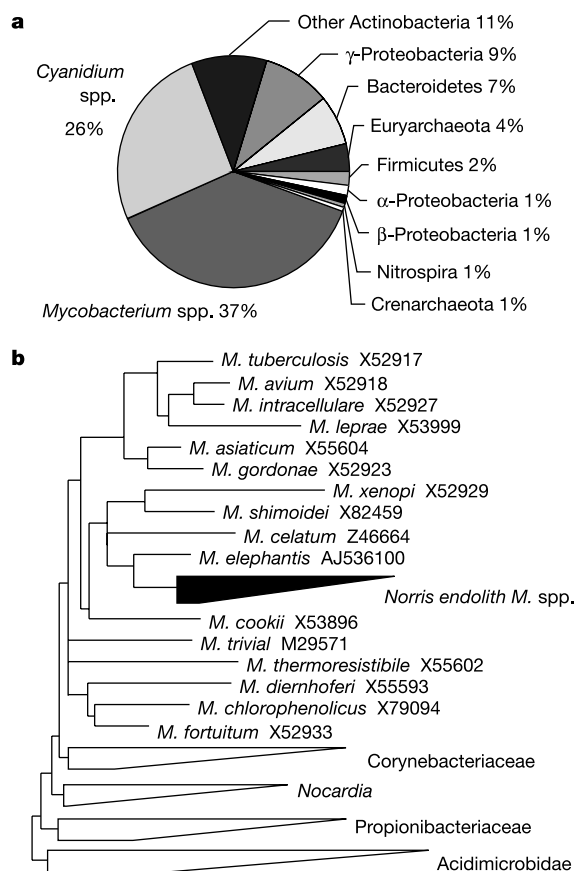


Figure 2 Microbial diversity of the Norris community. **a**, Phylogenetic distribution of rRNA gene sequences amplified from the Norris community ($n = 864$). **b**, Composite tree diagram of the phylogenetic relationship of endolithic community rRNA gene sequences (represented by solid wedge) to those of known *Mycobacterium* spp. (expanded portion of tree) and the relationship of the mycobacteria group to other groups (open wedges) of the bacterial division Actinobacteria. Phylogenetic analyses showed that rDNA sequences from the Norris community formed a distinct group of mycobacteria supported in all analyses (minimum evolution, maximum parsimony and maximum likelihood) with strong statistical support (>70% bootstrap value). The Norris sequences are most closely related to *M. elephantis* (~97%), and the association of the *M. elephantis* group as a sister species to the Norris group was present in all trees in all analyses and was supported statistically by maximum evolution bootstrap analysis but not by the other methods.

for lead, osmium and uranium, whereas the signal from the surrounding material indicated primarily silicon (Fig. 1k). Some cellular structures did not stain and had strong silicon signals, which indicates that they might already have been mineralized and could become fossilized.

The abundance of mycobacterial clones (~37%) detected in the community leads us to postulate that the ubiquitous curvilinear filamentous casts observed formed around a biofabric of primarily mycobacteria, although other organisms detected are probably also involved. Mycobacteria are known to form complex filamentous and branched biofilms²¹. The filaments in this study have a tubular construction with inner diameters of 0.8–3.0 µm (average 1.9 ± 0.5 µm, $n = 60$), typical of bacterial cells. BSE–SEM imaging showed that the filamentous structures are segmented, and EDS analysis indicates that the interiors contain biological material and that the casts are primarily silica (Fig. 1m). Similar filamentous structures observed in specimens collected from silica-depositing thermal springs in Yellowstone have been shown to be fossilized filament casts of microbes⁹.

Mineralization of endolithic communities associated with geothermal environments might lead to the preservation of identifiable biosignatures. Fossils of endolithic microbial communities recently discovered in Antarctic sandstone preserve exquisite detail of the internal structures of cells²². We expect that the abundance of Yellowstone endolithic communities in environments favourable to mineralization leads to their preservation in the geological record. Unequivocal microfossils from the 2.0-Gyr-old Gunflint formation (Ontario, Canada) were formed by silica mineralization in an environment interpreted as analogous to that in Yellowstone^{23,24}, although the exact nature of that environment remains uncertain. Regardless of the environment, Gunflint microfossils demonstrate the potential for microfossils to form in siliceous environments and to persist in the Earth's geological record for a minimum of 2 Gyr.

Mineralized endolithic communities such as those described here could be a highly diagnostic indicator of past life on Mars. Evidence of past geothermal activity on Mars includes volcanism and possible hydrothermal systems, which could be analogous to those of Yellowstone^{14,25}. Missions to Mars target such areas in the search for evidence of life, past or present, on Mars. The endolith is a critical habitat to explore in that search. The rich diversity of endolithic life in the geothermal environment of Yellowstone, combined with the potential for biosignature preservation, indicates that rocks associated with former hydrothermal systems might be the best hope for finding evidence of past life on the martian surface. □

Methods

Sample description

Samples were collected from outcrops located in Norris Geyser Basin, Yellowstone National Park, USA. The community described was collected from a single site (44°43.92' N, 110°42.63' W).

Chemical analysis

Rock pH was measured as described²⁶. Pore water was extracted by centrifugation for standard chemical analysis by inductively coupled plasma (ICP)–atomic absorption spectrometry, ICP–mass spectrometry and ion chromatography as described²⁰.

Molecular community analysis

Molecular phylogenetic community rRNA analysis was performed with DNA extracted from crushed rock samples as described²⁰. The average yield of extracted DNA was 3.1 mg of DNA per gram of crushed rock. PCR primers used in this study were 27F, 515F and 1391R. *In situ* hybridizations for mycobacterial species were performed as described²⁷.

Microscopy

Petrographic thin sections were prepared from resin-embedded samples and imaged with light and laser-scanning confocal microscopy (LSCM) as described²⁸. Conventional SEM specimens were fixed in 4% glutaraldehyde, critical-point dried and coated with gold¹³. BSE–SEM was performed as described^{22,29} with an FEI Quanta 600 SEM, BSE detector and Princeton Gamma Tech EDS system.

Photosynthetic pigment analysis

Photosynthetic pigments extracted from rock samples were analysed as described³⁰. Strains of *Cyanidium* spp. cultured from Norris Geyser Basin for comparison were obtained from the Culture Collection of Microorganisms from Extreme Environments (CCMEE) and included strains NC-5-Cl-1 CCMEE 5584 and C-Cl-1 CCMEE 5564.

Received 17 December 2004; accepted 14 February 2005; doi:10.1038/nature03447.

1. Golubic, S., Friedmann, E. I. & Schneider, J. The lithobiontic ecological niche, with special reference to microorganisms. *J. Sedim. Petrol.* **51**, 475–478 (1981).
2. Hofmann, B. A. & Farmer, J. D. Filamentous fabrics in low-temperature mineral assemblages: are they fossil biomarkers? Implications for the search for a subsurface fossil record on the early Earth and Mars. *Planet. Space Sci.* **48**, 1077–1086 (2000).
3. Friedmann, E. I. & Ocampo, R. Endolithic blue-green algae in dry valleys—primary producers in Antarctic desert ecosystem. *Science* **193**, 1247–1249 (1976).
4. Friedmann, E. I. Endolithic microorganisms in the Antarctic cold desert. *Science* **215**, 1045–1053 (1982).
5. Bell, R. A. Cryptendolithic algae of hot semiarid lands and deserts. *J. Phycol.* **29**, 133–139 (1993).
6. Gross, W., Kuver, J., Tischendorf, G., Bouchaala, N. & Busch, W. Cryptendolithic growth of the red alga *Galdieria sulphuraria* in volcanic areas. *Eur. J. Phycol.* **33**, 25–31 (1998).
7. Mosser, J. L., Misser, A. G. & Brock, T. D. Bacterial origin of sulfuric acid in geothermal habitats. *Science* **179**, 1323–1324 (1973).
8. White, D. E., Hutchinson, R. A. & Keith, T. E. C. *The Geology and Remarkable Thermal Activity of Norris Geyser Basin, Yellowstone National Park, Wyoming*. (US Geol. Surv. Professional Paper no. 1456, Washington DC, 1988).
9. Cady, S. L. & Farmer, J. D. in *Evolution of Hydrothermal Ecosystems on Earth (and Mars?)* (eds Bock, G. R. & Goode, J. A.) 150–173 (Wiley, New York, 1996).
10. Asada, R. & Tazaki, K. Silica biomineralization of unicellular microbes under strongly acidic conditions. *Can. Mineral.* **39**, 1–16 (2001).
11. Blank, C. E., Cady, S. L. & Pace, N. R. Microbial composition of near-boiling silica-depositing thermal springs throughout Yellowstone National Park. *Appl. Environ. Microbiol.* **68**, 5123–5135 (2002).
12. Zierenberg, R. A. & Schiffman, P. Microbial control of silver mineralization at a sea-floor hydrothermal site on the northern Gorda Ridge. *Nature* **348**, 155–157 (1990).
13. Cady, S. L., Farmer, J. D., Grotzinger, J. P., Schopf, J. W. & Steele, A. Morphological biosignatures and the search for life on Mars. *Astrobiology* **3**, 351–368 (2003).
14. Farmer, J. D. & Des Marais, D. J. Exploring for a record of ancient Martian life. *J. Geophys. Res. Planets* **104**, 26977–26995 (1999).
15. Banfield, J. F., Moreau, J. W., Chan, C. S., Welch, S. A. & Little, B. Mineralogical biosignatures and the search for life on Mars. *Astrobiology* **1**, 447–465 (2001).
16. Des Marais, D. J. & Walter, M. R. Astrobiology: Exploring the origins, evolution, and distribution of life in the Universe. *Annu. Rev. Ecol. Syst.* **30**, 397–420 (1999).
17. Pace, N. R. A molecular view of microbial diversity and the biosphere. *Science* **276**, 734–740 (1997).
18. Stackebrandt, E. & Goebel, B. M. Taxonomic note: A place for DNA–DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol.* **44**, 846–849 (1994).
19. Iivanainen, E. K., Martikainen, P. J., Raisanen, M. L. & Katila, M. L. Mycobacteria in boreal coniferous forest soils. *FEMS Microbiol. Ecol.* **23**, 325–332 (1997).
20. Spear, J. R., Walker, J. J., McCullem, T. M. & Pace, N. R. Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proc. Natl Acad. Sci. USA* **102**, 2555–2560 (2005).
21. Hall-Stoodley, L. & Lappin-Scott, H. Biofilm formation by the rapidly growing mycobacterial species *Mycobacterium fortuitum*. *FEMS Microbiol. Lett.* **168**, 77–84 (1998).
22. Ascaso, C. & Wierzbos, J. The search for biomarkers and microbial fossils in Antarctic rock microhabitats. *Geomicrobiol. J.* **20**, 439–450 (2003).
23. Schopf, J. W., Barghoorn, E. S., Maser, M. D. & Gordon, R. O. Electron microscopy of fossil bacteria two billion years old. *Science* **149**, 1365–1367 (1965).
24. Walter, M. R. A hot spring analog for the depositional environment of Precambrian iron formations of the Lake Superior Region. *Econ. Geol.* **67**, 965–980 (1972).
25. Christensen, P. R. *et al.* Morphology and composition of the surface of Mars: Mars Odyssey THEMIS results. *Science* **300**, 2056–2061 (2003).
26. Doemel, W. N. & Brock, T. D. pH of very acid soils. *Nature* **229**, 574 (1971).
27. St. Amand, A. L., Frank, D. N., De Groot, M. A. & Pace, N. R. Microscopic detection of *Mycobacterium avium* complex organisms in tissue with specific rRNA oligonucleotide probes. *J. Clin. Microbiol.* **4**(43), (in the press).
28. Papineau, D., Walker, J. J., Mojzsis, S. J. & Pace, N. R. Composition and structure of microbial communities from stromatolites of Hamelin Pool in Shark Bay, Western Australia. *Appl. Environ. Microbiol.* (in the press).
29. Wierzbos, J. & Ascaso, C. Application of backscattered electron imaging to the study of the lichen rock interface. *J. Microsc.* **175**, 54–59 (1994).
30. Frigaard, N. U., Larsen, K. L. & Cox, R. P. Spectrochromatography of photosynthetic pigments as a fingerprinting technique for microbial phototrophs. *FEMS Microbiol. Ecol.* **20**, 69–77 (1996).

Acknowledgements We thank K. Harris, D. Papineau, R. Ley, C. Rumbaitis-del Rio and the Pace laboratory members for discussion; H. Haesler, C. Hendrix and the staff of Yellowstone National Park for their assistance and discussion; J. Maresca and D. Bryant for high-performance liquid chromatography analysis of photosynthetic pigments; F. Luiszer and J. Drexler for chemical analysis and support for SEM; P. Boni and T. Giddings for sample preparation; H. Kleebe and G. Zito of the Colorado School of Mines for use of the BSE–SEM; and R. Castenholz and the CCMEE for cultures of *Cyanidium* spp. This research was supported by a grant from the NSF to N.R.P. and the University of Colorado Center for Astrobiology.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to N.R.P. (nrpace@colorado.edu). The GenBank database accession numbers for rRNA clone sequences are AY911422 to AY911496.

Species diversity can drive speciation

Brent C. Emerson & Niclas Kolm

Centre for Ecology, Evolution and Conservation, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK

A fundamental question in evolutionary ecology and conservation biology is: why do some areas contain greater species diversity than others? Island biogeographic theory has identified the roles of immigration and extinction in relation to area size and proximity to source areas^{1,2}, and the role of speciation is also recognized as an important factor^{3–6}. However, one as yet unexplored possibility is that species diversity itself might help to promote speciation, and indeed the central tenets of island biogeographic theory support such a prediction. Here we use data for plants and arthropods of the volcanic archipelagos of the Canary and Hawaiian Islands to address whether there is a positive relationship between species diversity and rate of diversification. Our index of diversification for each island is the proportion of species that are endemic, and we test our prediction that this increases with increasing species number. We show that even after controlling for several important physical features of islands, diversification is strongly related to species number.

Investigating why some areas contain greater species diversity than others offers a way to identify factors important for both generating and maintaining species diversity. The analysis of species–area curves has produced a wealth of data, with the evidence clearly supporting the simple rule that if you sample a larger area, you will find more species⁷. In their theory of island biogeography MacArthur and Wilson^{1,2} developed this basic principle to demonstrate that islands with non-zero immigration rates and different degrees of isolation should exhibit different degrees of species diversity. The consequences of this are demonstrated in numerous studies showing that in island systems, not only do larger islands contain more species, but for equivalent-sized islands those closer to a mainland source harbour a greater diversity of species. An implicit feature of MacArthur and Wilson's work is that both immigration rate and extinction rate are nonlinear. As the number of species on an island increases, extinction rate increases at a greater than linear rate. MacArthur and Wilson considered the increased interactions on an island with many species and reasoned that an increased number of species increases the likelihood of any given species dying out. So in the early stages of the colonization of an island and the

development of a community of species, the extinction rate would be low because there would be little competition. However, as more species arrived, competition would assume a greater role, with population sizes tending to become lower on average and species facing an increased probability of competitive exclusion and predation⁷.

An important but untested outcome of the situation outlined above is that the same factors that lead to an increased probability of extinction might also lead to an increased probability of diversification. In other words, species diversity itself could be a driver of species diversification. Increased competition and predation can have two consequences for any individual species on a given island. A species may succumb to these pressures and become extinct, or it may respond to these pressures, adapt, and survive. Additionally, as species diversity increases and average population size decreases, population genetic theory predicts that there will be an increasing probability of divergence of island populations from the source population of a species through genetic drift. Finally, increasing species diversity may lead to greater community structural complexity, and this has been suggested as a possible evolutionary force driving speciation⁸. To test the possibility that species diversity can drive speciation requires a demonstration that areas with greater species diversity also have greater diversification rates. Volcanic island archipelagos such as the Canary and Hawaiian Islands provide the best opportunity to examine this question because of their discrete geographical nature and the general absence of historical connection between their component islands to each other or to a mainland, although there have been historical connections between a few of the Hawaiian islands^{9,10}. Model systems such as the Canary and Hawaiian Islands permit the relative influences of variables to be readily considered compared with more complex continental areas¹¹. The flora and fauna of each island is the result of the accumulation of species through time by chance dispersal from mainland sources or neighbouring islands, *in situ* speciation and extinction. This contrasts with mainland biogeographic areas that historically may have had differing degrees of floral and faunal connectivity due to climate-mediated contractions and expansions in species range.

Both the Canary and Hawaiian Islands have an exceptionally well characterized flora and fauna for each of their component islands, and maximum geological ages are available for each island. To test our prediction that higher species diversity should lead to a higher rate of diversification, we investigated whether there is a positive relationship between the number of native species on an island and the diversification rate for arthropods and plants within each archipelago. By focusing on plants and arthropods, we were able to cover two highly diverse and disparate groups of organisms to test the generality of our hypothesis. Our index of diversification for individual islands is the proportion of native species that are endemic to that island. This includes species arising through intrinsic factors within the island (Fig. 1a, b) and those arising as a consequence of extrinsic factors resulting in extinction or speciation events on other islands (Fig. 1c, d). Thus our index is an estimate of diversification (speciation within individual clades averaged over all clades) within an island (Fig. 1a, b), with an inflation factor resulting from extinction and speciation on other islands (Fig. 1c, d). The impact of this inflation factor is greater for

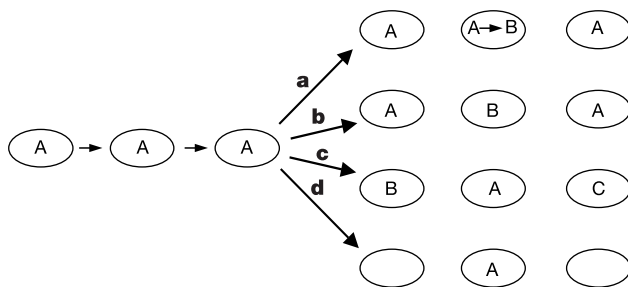


Figure 1 Origin of island endemism. Species A originates in the westernmost island and then colonizes the remaining two islands. After this there are four possible scenarios that may result in endemism on the middle island. **a**, Species B is the result of intra-island speciation from the founding species A. **b**, Species B is the result of inter-island speciation. **c**, Inter-island speciation on the first and third islands results in endemism for species A on the middle island. **d**, Extinction of species A on the first and third islands results in endemism on the middle island. In these last two scenarios, endemism of species A is not due to intrinsic factors within the island in which it occurs.

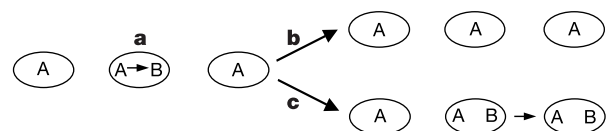


Figure 2 Elimination of island endemism. Intra-island speciation results in endemism of species B on the middle island (**a**). Either extinction (**b**) or the colonization of a new island (**c**) will eliminate this endemism.

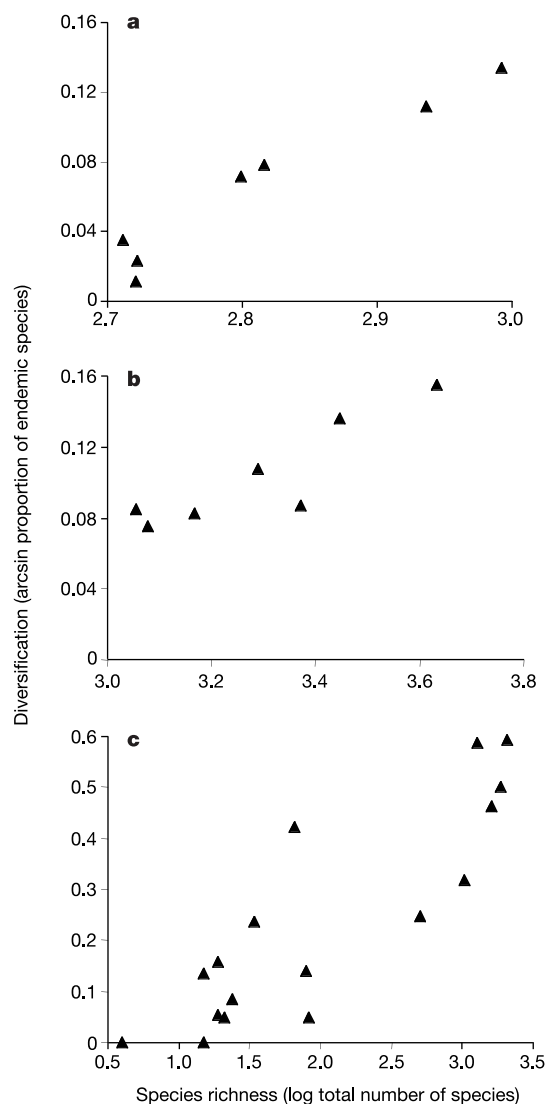


Figure 3 Bivariate plots for the relationship between diversification and species richness. **a**, Plants in the Canary Islands: $n = 7$, $r = 0.97$, $P = 0.0001$. **b**, Arthropods in the Canary Islands: $n = 7$, $r = 0.89$, $P = 0.007$. **c**, Arthropods in the Hawaiian Islands: $n = 17$, $r = 0.86$, $P = 0.0001$.

islands where we predict percentage endemism to be lower (see below), weakening rather than strengthening our predicted positive relationship between species diversity and diversification rate. Our index of diversification will also be diminished by intrinsic factors resulting in extinction (Fig. 2a) and extrinsic factors resulting in the

successful establishment of endemics onto new islands (Fig. 2b) leading to a reduction of endemic species on a given island. Here again, both these factors will act to weaken rather than strengthen our predicted relationship. Extinction will have more effect on islands where we predict percentage endemism to be highest (see below) and, considering the probability of successfully colonizing a new island to be equivalent among endemics, this will feature more for islands with a greater percentage of endemics.

To allow for the obvious biogeographic and compositional differences between the two archipelagos we performed separate analyses for each. Data for vascular plants and arthropods for the Canary Islands ($n = 7$: El Hierro, La Palma, La Gomera, Tenerife, Gran Canaria, Lanzarote, Fuerteventura) were extracted from refs 12 and 13, respectively. Data for flowering plants ($n = 18$: Kure, Midway, Pearl and Hermes, Lisianski, Laysan, Gardner, French Frigate Shoals, Necker, Nihoa, Kaula, Ni'ihau, Kaua'i, O'ahu, Moloka'i, Lana'i, Maui, Kaho'olawe, Hawai'i) and arthropods ($n = 17$: as for plants except for Kaula) of the Hawaiian Islands were extracted from refs 11 and 14, respectively. Each of the four lists of species is a complete inventory of all described presumed native species, including presumed extinct species and excluding anthropogenic introductions.

We used multiple regression analyses (see Methods) to control for other physical factors that might also be important for diversification rates within island ecosystems: island age, island area, island altitude, and proximity to the nearest neighbouring island. Older islands may be expected to contain proportionately more endemic species as a consequence of the greater time during which speciation could have occurred within them⁵. Larger and topographically more complex islands might also contain proportionately more endemic species because of an increased probability for intra-island allopatric speciation^{5,15}. Higher islands support more habitat diversity than lower ones¹⁶ and this could lead to proportionately more endemics. Islands closer to a neighbouring island might be expected to contain proportionately fewer endemic species because of a greater probability of receiving immigrants¹, increasing the proportion of non-endemics. Although these other physical factors may be considered particular to island ecosystems, our principal factor of interest, species diversity, is of much broader biological interest.

Our analyses show that species richness is indeed a strong predictor of the level of diversification both across taxa and island groups (Table 1). For arthropods on the Canary Islands, species richness was in fact the only independent variable that entered the final model and explained a vast amount of the variation in diversification (Table 1). Among the physical factors of the islands that we considered, island altitude covaried positively with diversification for plants on both the Hawaiian and Canary Islands and for arthropods on the Hawaiian Islands. Island area was a significant negative covariate with diversification for plants on the Canary

Table 1 Regression results (diversification as the dependent variable)

Island group	Taxon	Independent variables entering the final model	<i>n</i>	β	s.e. of β	<i>t</i>	<i>P</i>
Canary Islands	Arthropods	Number of species	7	0.89	0.20	4.45	0.007
		Number of species	7	1.02	0.07	14.3	0.0007
	Plants	Island altitude		0.10	0.06	1.7	0.18
		Island area		-0.22	0.05	4.4	0.02
Hawaiian Islands	Arthropods*	Number of species	17	0.74	0.10	7.3	< 0.0001
		Island altitude		0.54	0.12	4.6	0.0005
		Distance to nearest island		0.54	0.10	5.7	< 0.0001
	Plants†	Number of species	18	0.61	0.17	3.6	0.0028
		Island altitude		0.67	0.17	4.0	0.0013
		Distance to nearest island		0.70	0.18	3.8	0.0020

Separate forward stepwise multiple regression analyses were performed for each island group and taxon. Canary Islands, arthropods (final model): multiple $r^2 = 0.80$, $F_{(1,5)} = 19.8$, $P = 0.007$. Canary Islands, plants (final model): multiple $r^2 = 1.0$, $F_{(3,3)} = 241.2$, $P = 0.0004$. Hawaiian Islands, arthropods (final model): multiple $r^2 = 0.93$, $F_{(3,13)} = 60.7$, $P < 0.0001$. Hawaiian Islands, plants (final model): multiple $r^2 = 0.79$, $F_{(3,13)} = 17.6$, $P = 0.0005$. β is the standardized regression coefficient, *t* represents the test statistic.

* Because of collinearity between island altitude and island age, island age was dropped from the analysis of plants on the Hawaiian Islands (see Methods). Thus we cannot separate the effects of these two variables on diversification for this taxon and island group.

† Because of collinearity between number of species and island area, island area was dropped from the analysis of plants on the Hawaiian Islands (see Methods). Thus we cannot separate the effects of these two variables on diversification for this taxon and island group.

Islands, and distance to nearest island covaried positively with diversification for both plants and arthropods on the Hawaiian Islands. For plants on the Hawaiian Islands we could not separate the effects of species richness and island area because they were strongly correlated (see Methods). The strong relationship between diversification and species richness is revealed more readily with bivariate plots (Fig. 3) for vascular plants of the Canary Islands and arthropods of the Canary and Hawaiian Islands.

The strength of our analyses comes from the inclusive sampling of species within large taxonomic groups. The probability of speciation (and extinction) increases for each species on an island as total species number increases. However, this increase is small and is therefore unlikely to be apparent within an individual genus of flowering plants or arthropods. Of all factors considered in our analyses, species number was positively related to diversification for three of the multiple regression analyses, and was implicated in the fourth. Although other physical island factors also featured in the analyses, there was less consistency in both the number of comparisons in which they featured and the extent to which they offered positive explanatory power. Our results support an implicit but overlooked prediction of island biogeographic theory: as species number in an area increases, so should the rate of speciation. This is of great importance for the general understanding of patterns of species richness and community composition, and calls into question the validity of model-based approaches that do not incorporate this. The answer to questions such as why there are so many species in the tropics⁸ might in part be because there are so many species in the tropics. Further testing of the relationship between species number and rate of diversification can be achieved as complete faunal and floral lists for other island archipelagos come to hand. Recent developments in using microbial systems to test evolutionary ecological theory might also provide further experimental evidence^{17–19}. □

Methods

We used forward stepwise multiple regression analyses to control for the physical properties of the islands while investigating the relationship between diversification and species richness. All variables were transformed before analysis by using log transformation (species richness, island age, island area, island altitude, and proximity to the nearest neighbouring island) or arcsin transformation (diversification; that is, the proportion of endemic species) to ensure normality. Because collinearity between independent variables might confound the analyses we checked for redundancy by investigating tolerance levels for the variables for each separate analysis. Tolerance values were adequately high²⁰ for all except two pairs of variables across the four analyses (that is, more than 0.13), and no standard errors in β were inflated. Collinearity featured within the analyses of plants and arthropods for the Hawaiian Islands only. For plants, we detected a strong positive correlation between total number of species and island area ($r = 0.94$, $P = 0.0001$) and therefore excluded island area from the analysis. This means that we cannot separate the effect of number of species from the effect of island area on the diversification of flowering plants on the Hawaiian Islands. For arthropods, we detected a strong negative correlation between island altitude and island age ($r = -0.94$, $P < 0.0001$); island age was therefore removed from the analysis. This has no bearing on our result of the importance of species richness for this taxon and island group, although it means we cannot separate the effect of island altitude from the effect of island age on the diversification of arthropods on the Hawaiian Islands. To further assess that our results were not simply artefacts of our forward stepwise regression procedure²¹, we also performed standard multiple regressions for all four combinations of islands and taxa. This procedure produced similar results to those of the forward stepwise regressions (all models were significant, with total number of species as a significant explanatory variable for diversification) except for arthropods on the Canary Islands. For this analysis the full model was not significant because of the limited data set and the lack of power caused by the many independent variables. Our results from the stepwise multiple regression analyses were therefore not simply caused by random inflation of independent variables. We also investigated interaction effects between species richness and the four physical variables of the islands in the standard multiple regression analyses. We checked separately for each variable, but none of the interaction terms was significant within any of the analyses. Because our measure of diversification is a proportion of, and hence not completely independent of, the key variable in our analyses, species richness, we performed additional analyses using the same measure of diversification but with only non-endemic species as the measure of species richness (see Supplementary Information). Because the results did not change (Supplementary Table 2) using this procedure, it is highly unlikely that our results are confounded by non-independence between diversification and species richness.

Received 30 November 2004; accepted 10 February 2005; doi:10.1038/nature03450.

1. MacArthur, R. H. & Wilson, E. O. An equilibrium theory of insular zoogeography. *Evolution* **17**, 373–387 (1963).
2. MacArthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton Univ. Press, Princeton, 1967).
3. Gillespie, R. Community assembly through adaptive radiation in Hawaiian spiders. *Science* **303**, 356–359 (2004).
4. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, 2001).
5. Losos, J. B. & Schluter, D. Analysis of an evolutionary species–area relationship. *Nature* **408**, 847–850 (2000).
6. Heaney, L. R. Dynamic disequilibrium: a long-term, large-scale perspective on the equilibrium model of island biogeography. *Global Ecol. Biogeogr.* **9**, 59–74 (2000).
7. Rosenzweig, M. L. *Species Diversity in Time and Space* (Cambridge Univ. Press, Cambridge, 1995).
8. Tokeshi, M. *Species Coexistence: Ecological and Evolutionary Perspectives* (Blackwell, Oxford, 1999).
9. Carson, H. L. & Clague, D. A. in *Hawaiian Biogeography: Evolution on a Hotspot Archipelago* (eds Wagner, W. L. & Funk, V. A.) 14–29 (Smithsonian Institution Press, Washington DC, 1995).
10. Price, J. P. & Elliott-Fisk, D. L. Topographic history of the Maui Nui complex, Hawai'i, and its implications for biogeography. *Pacif. Sci.* **58**, 27–45 (2004).
11. Price, J. P. Floristic biogeography of the Hawaiian Islands: influences of area, environment and paleogeography. *J. Biogeogr.* **31**, 487–500 (2004).
12. Acebes Ginovés, J. R., et al. in *Lista de Especies Silvestres de Canarias (Hongos, Plantas y Animales Terrestres)* (eds Izquierdo, I., Martín, J. L., Zurita, N. & Arechavaleta, M.), 2nd edn (Consejería de Política Territorial y Medio Ambiente Gobierno de Canarias, La Laguna, Tenerife, in the press).
13. Oromi, P. & Báez, M. in *Lista de Especies Silvestres de Canarias (Hongos, Plantas y Animales Terrestres)* (eds Izquierdo, I., Martín, J. L., Zurita, N. & Arechavaleta, M.), 2nd edn (Consejería de Política Territorial y Medio Ambiente Gobierno de Canarias, La Laguna, Tenerife, in the press).
14. Nishida, G. M. Bishop Museum—Hawaiian arthropod checklist. (<http://www2.bishopmuseum.org/HBS/checklist/query.asp?grp=Arthropod>) (2002).
15. Roos, M. C., Kessler, P. J. A., Gradstein, S. R. & Baas, R. Species diversity and endemism of five major Malesian Islands: diversity–area relationships. *J. Biogeogr.* **31**, 1893–1908 (2004).
16. Hobohm, C. Plant species diversity and endemism on islands and archipelagos, with special reference to the Macaronesian Islands. *Flora* **195**, 9–24 (2000).
17. Buckling, A., Kassen, R., Bell, G. & Rainey, P. B. Disturbance and diversity in experimental microcosms. *Nature* **408**, 961–964 (2000).
18. Kassen, R., Buckling, A., Bell, G. & Rainey, P. B. Diversity peaks at intermediate productivity in a laboratory microcosm. *Nature* **406**, 508–512 (2000).
19. Rainey, P. B. & Travisano, M. Adaptive radiation in a heterogeneous environment. *Nature* **394**, 69–72 (1998).
20. Quinn, G. P. & Keough, M. J. *Experimental Design and Data Analysis for Biologists* (Cambridge Univ. Press, Cambridge, 2002).
21. Draper, N. R., Guttman, I. & Lapczak, L. Actual rejection levels in a certain stepwise test. *Commun. Statist. A* **8**, 99–105 (1979).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank W. Sutherland, G. Hewitt, M. Taylor and T. Givnish for constructive comments.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to B.C.E. (b.emerson@uea.ac.uk).

Sexual reproduction between partners of the same mating type in *Cryptococcus neoformans*

Xiaorong Lin, Christina M. Hull* & Joseph Heitman

Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina 27710, USA

* Present address: Departments of Biomolecular Chemistry, Medical Microbiology and Immunology, University of Wisconsin, Madison, Wisconsin 53706, USA

Cryptococcus neoformans is a globally distributed human fungal pathogen that causes life-threatening meningoencephalitis in immunocompromised patients¹. It has a defined sexual cycle involving haploid cells of α and a mating types², yet the vast majority of environmental and clinical isolates are α (ref. 3). Sexual recombination is normally expected to occur between

isolates of opposite mating type in organisms with two mating types (or sexes). How sexual reproductive potential can be maintained in an organism with a largely unisexual, nearly clonal population genetic structure is unknown. One clue, however, is that α strains undergo fruiting, a process that resembles sexual mating⁴ but is thought to be strictly mitotic and asexual. We report here that hallmarks of mating occur during fruiting, including diploidization and meiosis. Pheromone response pathway elements and the key meiotic regulator Dmc1 are required for efficient fruiting. Furthermore, fusion and meiosis can occur between non-isogenic α strains, enabling genetic exchange. These studies reveal how sexual reproduction can occur between partners of the same mating type. These findings have implications for the evolution of microbial pathogens, as well as for parthenogenesis, cell fusion events and transitions between self-fertilizing and outcrossing modes of reproduction observed in both fungi and other kingdoms.

Cryptococcus neoformans is a basidiomycetous fungus that grows as a budding yeast in culture and in the infected host. It undergoes a filamentous, dimorphic transition during the sexual cycle to produce spores, the suspected infectious agent⁵. When nutrients are limiting, α and a cells secrete pheromones that trigger cell–cell fusion. In contrast with the budding yeast *Saccharomyces cerevisiae*, the two parental nuclei congress but do not fuse, and the resulting dikaryon grows filamentously⁶. The tips of the filaments differentiate into basidia, where nuclear fusion and meiosis occur. Repeated mitosis results in multiple, haploid basidiospores that bud onto the basidium surface, forming four long chains^{2,6}.

Although this laboratory-defined sexual cycle of *C. neoformans* has been known for three decades, the environmental and clinical predominance of α strains has posed a conundrum³. If most of the population is limited to one mating type, how would a sexual cycle occur in nature? Haploid *C. neoformans* α strains can also undergo a developmental transition involving filamentation and sporulation,

known as haploid or monokaryotic fruiting⁴. This process resembles mating but was assumed to be mitotic and asexual. We show here that fruiting in fact represents a form of sexual reproduction between strains of the same mating type.

Fruiting and mating are both stimulated by similar environmental conditions (nitrogen starvation, desiccation, darkness and pheromones^{4,5,7}), and both involve hyphal growth and the production of basidia and spores. However, the two pathways do have distinguishing features; hyphal cells produced during fruiting are mononucleate, with unfused clamp connections, whereas mating hyphal cells are dikaryotic, linked by fused clamps.

Compared with mating, fruiting is inefficient, requires prolonged incubation, and occurs in a stochastic manner at isolated points at the periphery of a growth patch, suggesting that a rate-limiting step might restrict entry into this developmental pathway. The observation that diploid α/α strains undergo more rapid and robust fruiting compared with isogenic haploid α strains suggested that diploidization might normally occur during fruiting. Because isolating hyphal cells to determine ploidy is technically challenging, we made use of the fact that hyphae bud to produce vegetative yeast cells (termed blastospores)^{4,5}, which are readily isolated by micromanipulation and grow as budding yeast cells (Fig. 1). Blastospores are uninucleate, and their DNA content presumably reflects the nuclear content of the hyphae from which they are derived (by mitosis).

Haploid α cells (strain JEC21) were cultured on V8 medium for two weeks, and blastospores generated by the resulting hyphae were isolated by microdissection and subcultured. The vast majority (66/68, 97%) were diploid according to fluorescence-activated cell sorting (FACS) analysis (Fig. 1), uninucleate according to DNA staining, and of mating type α according to mating assays. Thus, we

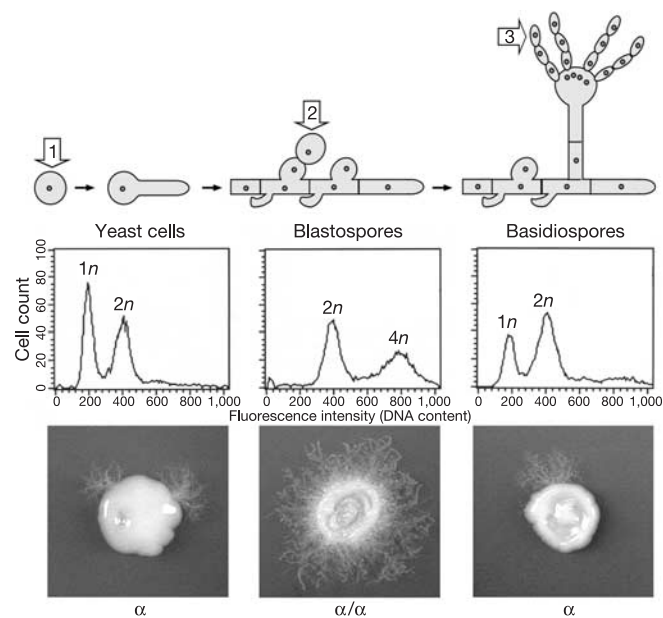


Figure 1 Diploidization during monokaryotic fruiting. The upper diagram depicts fruiting, which starts when haploid α yeast cells (1) filament in response to nutrient limitation and other signals. Large, round blastospores (2) are produced along the hyphae, whereas smaller, oval, basidiospores (3) are restricted to the basidium surface. Middle panels show fluorescence intensity (reflecting DNA content) of the original haploid α yeast cells, blastospore-derived α/α cells, and basidiospore-derived α cells. Lower panels show filamentation of the same three strains cultured for 3 weeks on filamentation agar in the dark.

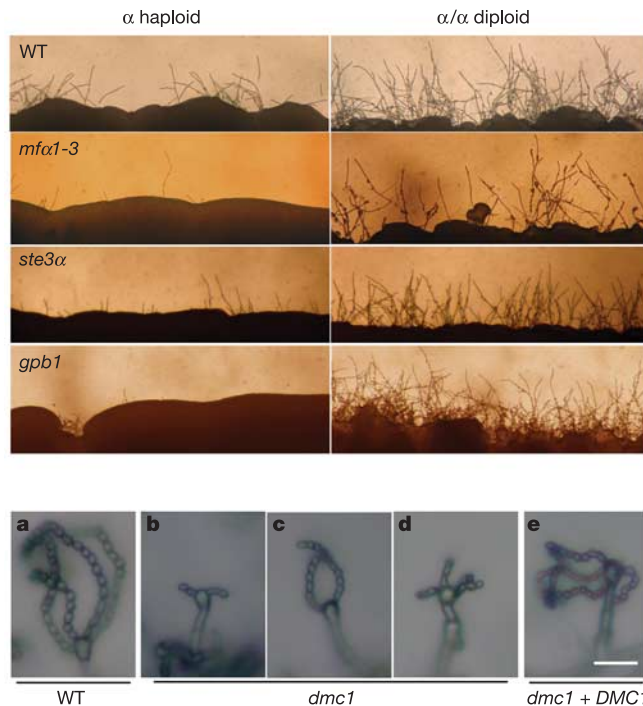


Figure 2 Pheromone response and meiotic components promote fruiting. Upper panels, filamentation assays of (left) the haploid wild-type (WT) α strain JEC21 and isogenic *mfx1-3*, *ste3 α* and *gpb1* mutants, and (right) the corresponding diploid wild-type and mutant α/α strains. **a–e**, Deletion of the meiosis-specific gene *DMC1* impairs sporulation during fruiting. **a**, A basidium with four long basidiospore chains produced by wild-type strain XL280. **b–d**, Basidia produced by *dmc1* mutant XL386 bearing two stunted (b), two intermediate length (c), or four short spore chains (d). **e**, A basidium with four long basidiospore chains produced by the *dmc1+DMC1* complemented strain. Scale bar, 10 μ m.

conclude that nuclear diploidization occurs during monokaryotic fruiting. This does not appear to be the result of promiscuous cell–cell or nuclear fusion, or a misregulated cell cycle, because ploidies greater than diploid were not observed (see Supplementary Information). All of the blastospore-derived diploid α/α strains undergo fruiting more robustly than the haploid α parental strain (Fig. 1). Diploidization might directly promote filamentation, or the two processes might be different rate-limiting steps. This ploidy increase during monokaryotic fruiting might be conceptually similar to endoduplication in oocytes, which gives rise to parthogenetic offspring in insects and other multicellular eukaryotes.

Generation of α/α diploids during monokaryotic fruiting was found to depend on elements required for α and **a** cell fusion during mating. *C. neoformans* mating is controlled by pheromones and pheromone receptors, G proteins and a MAP kinase pathway⁵. α haploid strains lacking the three MF α pheromones, the pheromone receptor Ste3 α , or the pheromone-activated G β subunit Gpb1 showed a defect in fruiting and produced reduced filaments^{7,8} (Fig. 2a). Blastospores were recovered from these mutants and subjected to FACS analysis; the proportion of α/α diploids to α haploids was reduced 5- to 15-fold compared with wild-type α cells (*mfa1-3 Δ* , *ste3 α Δ* and *gpb1 Δ* mutants produced 7% (2/27), 8% (5/64) and 21% (6/29) α/α blastospores, respectively). Re-introduction of the wild-type genes restored wild-type fruiting behaviour (not shown). When these same mutations were analysed in α/α homozygous mutants, none of the gene products was required, and filamentation occurred at α/α cell wild-type levels (Fig. 2a). We conclude that the pheromone response pathway functions before cell or nuclear fusion during fruiting. In *S. cerevisiae*, pheromone actively promotes cell and nuclear fusion during mating, and α -factor is required for nuclear fusion in **a**–**a** dikaryons produced by protoplast fusion⁹. Pheromone signalling might play an analogous role in *C. neoformans* fruiting.

We next addressed whether nuclear fusion first requires cell–cell fusion. Using α strains marked by transformation with dominant drug resistance markers (nourseothricin (NAT) or neomycin (NEO)), we determined the frequency of doubly resistant diploid progeny (confirmed by FACS) following co-culture on V8 medium. The frequency of diploid progeny ranged from 0.1 to 14 per million colony-forming units (c.f.u.). This very low frequency of spontaneous cell fusion suggests that another mechanism, such as

endoduplication or stimulated cell fusion, operates during fruiting. In fact, previous studies have established that **a** pheromone stimulates fruiting of α strains⁷ (α pheromone does so to a lesser extent) and **a** cells trigger α – α cell fusion¹⁰. We found that fusion between auxotrophic α strains (XL341 α *ura5 lys1* and XL342 α *ade2*) occurs at a low frequency but is stimulated ~1,000-fold by the presence of counter-selectable **a**-cells (JEC169 **a** *ura5 lys1 ade2*). This might represent a mechanism by which α cells in nature could be stimulated to fuse with one another when conditions are not conducive to traditional mating (for example, the presence of few or distant **a** cells).

Because fruiting involves diploidization and then reduction to haploid basidiospores, we tested whether chromosome assortment and recombination also occur. A diploid α/α strain (CHY601, *lys1/LYS1 lys2/LYS2 ura5/URA5 ade2/ADE2*) was induced to undergo monokaryotic fruiting, and basidiospores were isolated on 5-fluoroorotic acid (5-FOA) medium to select *ura5* segregants. Of 257 basidiospore isolates analysed, all mated as α cells, and the auxotrophic markers that were heterozygous in the parental strain showed mendelian segregation (see Supplementary Information). All of the progeny analysed (20/20) were haploid by FACS analysis, with no evidence for aneuploidy. In control experiments, no haploidization of the α/α diploid strain (CHY601) was observed among 5-FOA-resistant isolates. Thus, during fruiting, diploid α/α cells undergo independent chromosome assortment and a complete genome reduction to produce haploid progeny.

A high rate of recombination was also found to occur during fruiting. Independent chromosome assortment could occur through sexual or parasexual reproduction. In parasexual life cycles, ploidy is reduced by random chromosome loss with no increased recombination. In contrast, genetic recombination is frequent in sexual reproduction. We analysed recombination during fruiting using a recently developed meiotic map¹¹. An α/α diploid strain was constructed with two F₁ progeny (XL304 α NEO fused with XL202 α NAT) derived from a cross between the mapped strains. Basidiospores produced by fruiting from this α/α diploid were isolated by microdissection. All were α and haploid (on the basis of mating tests and FACS analysis), and the NAT and NEO transgenes showed independent assortment (Table 1). By restriction-fragment length polymorphism (RFLP) analysis of 20 markers from 6 linkage groups on 5 chromosomes, recombination was found to occur across the

Table 1 Recombination occurs frequently during monokaryotic fruiting

CHR*	Linkage group	RFLP marker	Distance† (cM)	Progeny								XL304 (P1)	XL202 (P2)
				1 NEO	2 NAT NEO	3 NAT NEO	4	5 NEO	6 NAT NEO	7 NEO	8		
1	12	<i>Hind32</i>	3.3	0	0	0	0	1	1	0	1	0	1
1	12	<i>Xho10</i>	70.9	0	0	0	1	1	0	1	1	0	1
1	12	<i>Pst9</i>	84.0	0	1	0	1	1	0	1	1	0	1
3	6	<i>Xba13</i>	0	1	0	0	1	1	1	1	0	0	1
3	6	<i>Eco4</i>	48.6	1	0	0	1	1	1	1	0	0	1
5	7	<i>Pst15</i>	16.7	0	0	1	0	0	0	1	1	0	1
5	7	<i>Eco24</i>	44.2	0	0	1	0	0	0	1	0	0	1
5	7	<i>Eco13</i>	102.8	1	0	1	0	0	0	1	0	0	1
5	16	<i>Eco26</i>	0	1	1	1	1	0	0	1	1	0	1
5	16	<i>Bam14</i>	11.2	1	1	1	0	1	0	1	0	0	1
5	16	<i>Hind33</i>	27.6	1	1	0	0	1	1	1	0	0	1
7	11	<i>Hind16</i>	64.6	1	1	0	1	1	1	0	1	0	1
7	11	<i>Pst14</i>	69.8	1	1	0	1	0	1	0	1	0	1
7	11	<i>Bam12</i>	72.9	1	1	1	0	0	1	0	1	0	1
7	11	<i>Eco27</i>	99.3	1	0	1	1	0	1	0	0	0	1
7	11	<i>Hind22</i>	99.7	1	0	1	1	0	1	0	0	0	1
7	11	<i>Eco2</i>	101.6	1	0	1	1	0	1	0	0	0	1
11	1	<i>Bam13</i>	0	1	1	0	0	0	1	1	0	0	1
11	1	<i>Xho23</i>	16.1	1	1	0	1	0	1	1	0	0	1
11	1	<i>Eco3</i>	157.4	1	0	1	0	1	1	1	1	0	1

The diploid strain XL202/XL304 (α/α NAT NEO) was cultured on V8 medium for three weeks. Progeny derived from dissected basidiospores, which were all α and haploid (on the basis of mating-type assays and FACS analysis) were scored for RFLP markers heterozygous in the parental diploid strain, based on the meiotic map of Marra *et al.*¹¹. Bold text indicates recombination events. P1 and P2 denote the two parental strains.

* CHR, chromosome number.

† Genetic distance of the markers from the end of the linkage group, on the basis of the meiotic map.

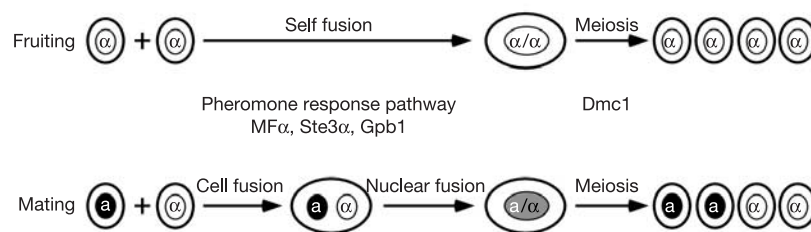


Figure 3 Sexual recombination through mating or fruiting. During fruiting, α cells diploidize, filament, produce basidia, undergo meiosis and sporulate, producing recombinant haploid progeny. During mating, nuclear fusion occurs in the basidium,

where meiosis and sporulation also occur. Pheromone response pathway elements and the meiotic regulator Dmc1 control both processes.

majority of chromosomes examined (5 of 6 linkage groups) at a rate similar to a - α cell sporulation (Table 1). All progeny showed unique genotypes, none identical to either parent, revealing that a high frequency of recombination occurs throughout the genome during fruiting, consistent with meiotic recombination.

Fruiting was also found to require a key meiotic regulator. The *DMC1* gene, the conserved homologues of which mediate DNA double-strand-break repair and homologous chromosome pairing during meiosis in yeasts and mammals¹², was found to be essential for sporulation during mating (α *dmc1* $\times$ a *dmc1*) (not shown) and to play a critical role in sporulation and spore viability during fruiting (α *dmc1*). The α wild-type strain underwent fruiting (Fig. 2b), and 100% of the resulting basidia (35/35) bore four spore chains with an average of 13.5 ± 3.5 spores per chain. The germination frequency of dissected spores was 27.3% (24/88). In contrast, the α *dmc1* mutant strain underwent fruiting to produce 48% (30/62) of basidia with four spore chains and 52% (30/62) with only two spore chains (dyads) (Fig. 2b). Moreover, the spore chains produced were $\sim 50\%$ shorter than wild type (7.3 ± 6 spores per chain), and the vast majority of spores failed to germinate (germination frequency 2/133 or 1.5%). Sporulation defects conferred by the *dmc1* mutation were complemented by the wild-type *DMC1* gene (Fig. 2). Thus, mutation of the meiosis-specific *DMC1* gene severely impairs sporulation and spore viability during fruiting, indicating that meiotic recombination plays an important role.

Our findings reveal that *C. neoformans* can reproduce sexually by either mating or monokaryotic fruiting. The two processes share several key features, including cell fusion, filamentous growth, basidia formation, meiosis and sporulation (Fig. 3). Key distinguishing features are that nuclear fusion occurs early during fruiting, and late during mating, and that fruiting involves partners of the same mating type whereas mating is restricted to those of opposite mating type. The discovery that meiosis can occur in α/α diploids of *C. neoformans* is analogous to *S. cerevisiae* *rme1* mutants. Meiosis in *S. cerevisiae* is normally restricted to a/α cells owing to repression of the meiotic inhibitor Rme1 by the $a1/\alpha2$ complex¹³. In *rme1/rme1* mutants, however, meiosis is liberated from these constraints and both a/a and α/α strains can undergo meiosis and sporulate. Although no direct Rme1 homologue is apparent in the *C. neoformans* genome, repression of an analogous factor might govern fruiting of α/α strains.

Another fungal human pathogen, *Pneumocystis carinii*, also appears to undergo a haploid-diploid ploidy shift in the trophic stage¹⁴, and the cystic form contains 2, 4, or 8 nuclei in which synaptonemal complexes have been reported^{15,16}. These findings suggest that *P. carinii* might undergo a modified form of sexual reproduction similar to that of *C. neoformans*. Like *C. neoformans*, there is no evidence of mating-type switching in *P. carinii*, and only one mating type is known¹⁷. These similarities suggest that two divergent human fungal pathogens might make use of a modified sexual cycle (linked to infection and virulence) that involves cells of the same mating type. Endoduplication or mating between cells of

the same mating type can confer evolutionary advantages when the other mating type is rare, resulting in a change from a heterothallic (outcrossing) to a homothallic lifecycle. Similar changes from obligate outcrossing to self-fertilization occur as a prominent evolutionary transition in flowering plants. One mechanism that operates in *Arabidopsis thaliana* involves inactivation and positive selection of pseudogenes at the self-incompatibility S locus¹⁸, enabling the rapid evolution of self-pollination after loss of self-incompatibility. In *C. neoformans*, a similar change could have evolved either in response to distortion of the two mating types (sexes) in the environment, or have provided the means for this disparity to have arisen.

In many pathogenic fungi, including *C. neoformans* and *Candida albicans*, the population structure is largely clonal, yet the organisms maintain their ability to undergo mating infrequently, potentially to allow response to environmental and host challenges¹⁹. In the human pathogenic parasite *Toxoplasma gondii*, sexual recombination is infrequent but is thought to have given rise to oral transmission and can enhance virulence^{20,21}. Thus, an evolutionary compromise between clonal and sexually recombining life cycles has been struck, and we hypothesize that mating between cells of the same mating type might contribute to both.

Given the preponderance of α strains in the environment, whether and how *C. neoformans* maintains sexual potential has been an important but unresolved question. The discovery that *C. neoformans* α cells can sexually reproduce via fruiting, without fusing with a partner of opposite mating type, provides a mechanism for a long-term survival advantage for α cells. For example, the ability to produce α/α diploid cells may enhance adaptation to changing environments. In *S. cerevisiae*, population genetic studies on drug resistance evolution reveal that haploid and diploid populations are each uniquely better at adapting to different environments²². Over the long term, a mixture of haploid and diploid cells may thus provide an even greater range of adaptability than either alone. Ploidy also influences gene expression and adaptive mutation fixation rates^{23–25}. Monokaryotic fruiting could be occurring through self-diploidization, or cell–cell fusion of genetically distinct partners of like mating type. In the case of self-diploidization, the benefits of meiosis might be to remove transposable elements, enable large-scale genome rearrangement, or to promote repair in mutagenic environments (which could include the host). In mice, multipotent hematopoietic cell fusion with hepatocytes gives rise to a unique cell population with increased ploidy that mediates hepatic regeneration. This example illustrates how fusion between cells with identical genomes but in different epigenetic states can alter developmental fate²⁶. By analogy, fungal fruiting involving endoduplication or cell–cell fusion may similarly provide access to new cell fates. By providing α cells access to a novel type of sexual cycle that can produce first diploids and then recombinant progeny, monokaryotic fruiting might have important roles in the ecology and evolution of *C. neoformans*, which can now be tested further under defined conditions. □

Methods

Strains

Strains used in this study were all derived from an inbred strain series, including the congenic strains JEC21 (*MATα*) and JEC20 (*MATa*) and related strains B3501 (*MATα*) and B3502 (*MATa*)²⁷ and JEC20/21 congenic derivatives JEC38 (*MATa met1*), JEC169 (*MATa ade2 lys1 ura5*), JEC170 (*MATα ade2 lys2*), JEC171 (*MATa ade2 lys2*), CHY601 (*MATα/α ade2/ADE2 lys1/LYS1 lys2/LYS2 ura5/URA5*), CHY604 (*MATα/α ade2/ADE2 lys1/LYS1 ura5/ura5*), CHY607 (*MATa/a ade2/ADE2 lys2/LYS2 ura5/ura5*), WSC40 (*MATα ade2 ste3α::ADE2*), WSC55 (*MATα mfx1::ADE2 mfx2,3::URA5 ade2 ura5*), WSC57 (*MATα mfx1::ADE2 mfx2,3::URA5 ade2 ura5 MFα1::URA5*), WSC72 (*MATα ura5 STE3::URA5 ade2 ste3α::ADE2*), WSC129 (*MATα ura5 gpb1::URA5*, ref. 8) and XL854 (*MATα GPB1 NAT ura5 gpb1::URA5*). Strains XL341 (*MATα ura5 lys1*) and XL342 (*MATα ade2*) are meiotic progeny dissected from the cross between strains JEC21 and JEC169 for this study. Strain XL854 was obtained by introduction of polymerase chain reaction (PCR) products of the wild-type copy of the *GPB1* gene with a linked NAT marker into mutant strain WSC129 by biolistic transformation. Strains XL366 (*MATα dmc1::NAT*) and XL367 (*MATa dmc1::NAT*) in the JEC21 genetic background, and strain XL386 (*MATα dmc1::NEO*) in the XL280 genetic background (described below) were constructed for this study (see Supplementary Information for details). Complemented strains XL886 and XL889 were obtained by introduction of PCR products of the wild-type copy of the *DMC1* gene with the NAT or NEO marker into mutant strains XL386 and XL366, respectively, by biolistic transformation.

Diploid α/α strains were constructed by assisted matings with a counter-selectable pheromone donor as previously described¹⁰, isolation and analysis of blastospore-derived colonies, or selection for fusion between strains marked with different dominant markers. To obtain diploid (*MATα/α*) isolates of the *mfx1-3*, *ste3α*, and *gpb1* mutants, strains RDC46-3 (*MATα mfx1-3*, ref. 8), WSC40 (*MATα ste3αΔ*) and WSC129 (*MATα gpb1Δ*) were cultured on V8 medium (pH 7.0) for three weeks and blastospores were dissected. Diploid isolates of the corresponding mutants XL894 (*MATα/α mfx1-3/mfx1-3*), XL518 (*MATα/α ste3αΔ/ste3αΔ*), XL520 (*MATα/α gpb1Δ/gpb1Δ*) derived from these blastospores were confirmed by FACS analysis. Strains XL202 (*MATα*), XL304 (*MATα*) and XL280 (*MATα*) are meiotic progeny dissected from the cross between strains B3501 (*MATα*) and B3502 (*MATa*) for this study. The diploid strain XL202/XL304 (α/α NAT NEO) was isolated from the co-culture of strains XL202 (α NAT) and XL304 (α NEO) on V8 juice agar medium and doubly resistant strains were selected on YPD agar medium containing both NAT and G418.

Microscopy

Cells were grown on V8 medium on slides in the dark. Hyphae were fixed in 3.7% formaldehyde in PBS with 1% Triton. Nuclei and septa were visualized by staining with 4',6'-diamidino-2-phenylindole (DAPI; Sigma) and Calcofluor (Sigma) as described previously⁴.

Genetic manipulations

Crosses were performed on V8 juice agar medium (pH 7.0) in the dark at 22 °C. Individual basidiospores or blastospores were dissected by micromanipulation and transferred to YPD agar plates. From these subcultures, progeny clones were isolated for analysis.

Mating type assays

To determine mating type, each isolate was separately cocultured with the parental tester strains JEC21 (*MATα*) and JEC20 (*MATa*) on V8 medium. Mating type was also determined by PCR using *STE20* gene primers that yield mating-type-specific amplicons.

Flow cytometry

Cells were processed for flow cytometry essentially as previously described²⁸ but with minor modifications. Briefly, cells were harvested after growth on YPD medium, washed in PBS buffer and fixed in 70% ethanol overnight at 4 °C. Fixed cells were washed in NS buffer (10 mM Tris-HCl pH 7.6, 250 mM sucrose, 1 mM EDTA (pH 8.0), 1 mM MgCl₂, 0.1 mM CaCl₂, 0.1 mM ZnCl₂, 0.4 mM phenylmethyl sulphonyl fluoride, 7 mM β -mercaptoethanol), stained with propidium iodide in NS buffer containing RNaseA at 4 °C for 4–16 h. Stained cells were diluted into Tris buffer and sonicated. Flow cytometry was performed on 10,000 cells and analysed using the FL1 channel on a Becton–Dickinson FACSscan.

Transformations

Dominant selectable markers conferring resistance to nourseothricin (NAT) and Geneticin 418 (NEO) markers²⁹ were transformed biolistically using a Bio-Rad Model PDS-1000/He Biolistic Particle Delivery System as previously described³⁰.

DMC1 gene disruption and reconstitution

The *DMC1* gene was deleted and replaced with the NEO marker in strain XL280 and with the NAT marker in strain JEC21 using overlap PCR and biolistic transformation, and confirmed by Southern blot. The wild-type *DMC1* gene linked to the NAT or G418 drug resistance marker was reintroduced for complementation studies. For details, see Supplementary Information.

Construction of α/α strain and recombinational analysis

The meiotic map is based on two F₁ sibling strains, B-3501 and B-3502, which are identical in 50% of the genome but contain DNA polymorphisms in the other 50% (ref. 11). Two hyper-filamentous α progeny polymorphic in approximately 25% of their genome from a cross between B3501 and B3502 (isolates XL304 and XL202) were marked with dominant

selectable markers and co-incubated on V8 medium. An α/α double resistant diploid was then selected on YPD medium using both NAT and G418. The resulting heterozygous α/α strain was fruited and DNA was prepared from the progeny and analysed by RFLP analysis of defined DNA polymorphisms as described¹¹.

Received 24 November 2004; accepted 9 February 2005; doi:10.1038/nature03448.

- Casadevall, A. & Perfect, J. R. *Cryptococcus neoformans* (ASM Press, Washington, DC, 1998).
- Kwon-Chung, K. J. Morphogenesis of *Filobasidiella neoformans*, the sexual state of *Cryptococcus neoformans*. *Mycologia* **68**, 821–833 (1976).
- Kwon-Chung, K. J. & Bennett, J. E. Distribution of α and a mating types of *Cryptococcus neoformans* among natural and clinical isolates. *Am. J. Epidemiol.* **108**, 337–340 (1978).
- Wickes, B. L., Mayorga, M. E., Edman, U. & Edman, J. C. Dimorphism and haploid fruiting in *Cryptococcus neoformans*: association with the α -mating type. *Proc. Natl Acad. Sci. USA* **93**, 7327–7331 (1996).
- Hull, C. M. & Heitman, J. Genetics of *Cryptococcus neoformans*. *Annu. Rev. Genet.* **36**, 557–615 (2002).
- McClelland, C. M., Chang, Y. C., Varma, A. & Kwon-Chung, K. J. Uniqueness of the mating system in *Cryptococcus neoformans*. *Trends Microbiol.* **12**, 208–212 (2004).
- Wang, P., Perfect, J. R. & Heitman, J. The G-protein β subunit GPB1 is required for mating and haploid fruiting in *Cryptococcus neoformans*. *Mol. Cell. Biol.* **20**, 352–362 (2000).
- Shen, W. C., Davidson, R. C., Cox, G. M. & Heitman, J. Pheromones stimulate mating and differentiation via paracrine and autocrine signaling in *Cryptococcus neoformans*. *Eukaryot. Cell* **1**, 366–377 (2002).
- Rose, M. D., Price, B. R. & Fink, G. R. *Saccharomyces cerevisiae* nuclear fusion requires prior activation by alpha factor. *Mol. Cell. Biol.* **6**, 3490–3497 (1986).
- Hull, C. M., Davidson, R. C. & Heitman, J. Cell identity and sexual development in *Cryptococcus neoformans* are controlled by the mating-type-specific homeodomain protein Sxi α . *Genes Dev.* **16**, 3046–3060 (2002).
- Marra, R. E. *et al.* A genetic linkage map of *Cryptococcus neoformans* variety *neoformans* serotype D (*Filobasidiella neoformans*). *Genetics* **167**, 619–631 (2004).
- Sauvageau, S., Ploquin, M. & Masson, J. Y. Exploring the multiple facets of the meiotic recombinase Dmc1. *Bioessays* **26**, 1151–1155 (2004).
- Mitchell, A. P. & Herskowitz, I. Activation of meiosis and sporulation by repression of the RME1 product in yeast. *Nature* **319**, 738–742 (1986).
- Thomas, C. F. Jr & Limper, A. H. Pneumocystis pneumonia. *N. Engl. J. Med.* **350**, 2487–2498 (2004).
- Matsumoto, Y. & Yoshida, Y. Sporogony in *Pneumocystis carinii*: synaptonemal complexes and meiotic nuclear divisions observed in precysts. *J. Protozool.* **31**, 420–428 (1984).
- Wyder, M. A., Rasch, E. M. & Kaneshiro, E. S. Quantitation of absolute *Pneumocystis carinii* nuclear DNA content. Trophic and cystic forms isolated from infected rat lungs are haploid organisms. *J. Eukaryot. Microbiol.* **45**, 233–239 (1998).
- Smulian, A. G., Sesterhenn, T., Tanaka, R. & Cushion, M. T. The *ste3* pheromone receptor gene of *Pneumocystis carinii* is surrounded by a cluster of signal transduction genes. *Genetics* **157**, 991–1002 (2001).
- Shimizu, K. K. *et al.* Darwinian selection on a selfing locus. *Science* **306**, 2081–2084 (2004).
- Johnson, A. The biology of mating in *Candida albicans*. *Nature Rev. Microbiol.* **1**, 106–116 (2003).
- Grigg, M. E., Bonnefoy, S., Hehl, A. B., Suzuki, Y. & Boothroyd, J. C. Success and virulence in *Toxoplasma* as the result of sexual recombination between two distinct ancestries. *Science* **294**, 161–165 (2001).
- Su, C. *et al.* Recent expansion of *Toxoplasma* through enhanced oral transmission. *Science* **299**, 414–416 (2003).
- Anderson, J. B., Sirjusingh, C. & Ricker, N. Haploidy, diploidy and evolution of antifungal drug resistance in *Saccharomyces cerevisiae*. *Genetics* **168**, 1915–1923 (2004).
- Galitski, T., Saldanha, A. J., Styles, C. A., Lander, E. S. & Fink, G. R. Ploidy regulation of gene expression. *Science* **285**, 251–254 (1999).
- Zeyl, C., Vanderford, T. & Carter, M. An evolutionary advantage of haploidy in large yeast populations. *Science* **299**, 555–558 (2003).
- Zeyl, C. Experimental studies on ploidy evolution in yeast. *FEMS Microbiol. Lett.* **233**, 187–192 (2004).
- Vassilopoulos, G., Wang, P. R. & Russell, D. W. Transplanted bone marrow regenerates liver by cell fusion. *Nature* **422**, 901–904 (2003).
- Kwon-Chung, K. J., Edman, J. C. & Wickes, B. L. Genetic association of mating types and virulence in *Cryptococcus neoformans*. *Infect. Immun.* **60**, 602–605 (1992).
- Sia, R. A., Lengeler, K. B. & Heitman, J. Diploid strains of the pathogenic basidiomycete *Cryptococcus neoformans* are thermally dimorphic. *Fungal Genet. Biol.* **29**, 153–163 (2000).
- Fraser, J. A., Subaran, R. L., Nichols, C. B. & Heitman, J. Recapitulation of the sexual cycle of the primary fungal pathogen *Cryptococcus neoformans* var. *gattii*: implications for an outbreak on Vancouver Island, Canada. *Eukaryot. Cell* **2**, 1036–1045 (2003).
- Toffaletti, D. L., Rude, T. H., Johnston, S. A., Durack, D. T. & Perfect, J. R. Gene transfer in *Cryptococcus neoformans* by use of biolistic delivery of DNA. *J. Bacteriol.* **175**, 1405–1411 (1993).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank W.-C. Shen for strains, A. Mitchell, J. Fraser and A. Idnurm for discussions, and T. Mitchell, D. Lew, B. Capel, R. Wharton and J. Anderson for critical reading. This work was supported by an NIAID R01 grant to J.H. C.M.H. was supported by a Damon Runyon Cancer Research Fellowship and a Burroughs Wellcome Career Development Award in the Biomedical Sciences. J.H. is a Burroughs Wellcome Fund Scholar in Molecular Pathogenic Mycology and an Investigator of the Howard Hughes Medical Institute.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.H. (heitm001@duke.edu) or X.L. (xrlin@duke.edu).

Regulation of axon growth *in vivo* by activity-based competition

Jackie Yuanyuan Hua^{1,2}, Matthew C. Smear³, Herwig Baier³ & Stephen J. Smith¹

¹Department of Molecular and Cellular Physiology, and ²Neurosciences Program, Stanford University, Stanford, California 94305, USA

³University of California, San Francisco, Department of Physiology, Program in Neuroscience, 1550 4th Street, San Francisco, California 94143-2722, USA

The formation of functional neural networks requires precise regulation of the growth and branching of the terminal arbors of axons, processes known to be influenced by early network electrical activity^{1–3}. Here we show that a rule of activity-based competition between neighbouring axons appears to govern the growth and branching of retinal ganglion cell (RGC) axon arbors in the developing optic tectum of zebrafish. Mosaic expression of an exogenous potassium channel or a dominant-negative SNARE protein was used to suppress electrical or neurosecretory activity in subsets of RGC axons. Imaging *in vivo* showed that these forms of activity suppression strongly inhibit both net growth and the formation of new branches by individually transfected RGC axon arbors. The inhibition is relieved when the activity of nearby ‘competing’ RGC axons is also suppressed. These results therefore identify a new form of activity-based competition rule that might be a key regulator of axon growth and branch initiation.

Pioneering studies of the development of mammalian peripheral motor axons have shown that axon arbor growth is governed by activity-based competition rules that appear to operate by means of an effect on branch elimination^{4,5}. Axon branches more active in releasing neurotransmitters persist at specific neuromuscular sites, whereas less active axon branches retract, resulting in the canonical elimination of polyneuronal innervation. There have been suggestions of analogous effects during the development of visual cortex^{6–8}, olfactory bulb^{9,10} and *Xenopus* optic tectum¹¹. Here we exploit the technically advantageous zebrafish optic tectum to test for a competition rule operating *in vivo* in a developing vertebrate brain. Previous work has firmly established a requirement of neural

activity in the refinement of the RGC tectal arbor and has identified some molecular pathways involved^{12–15}. We now report effects of suppressing activity in single RGC axons, with evidence for activity-based competition. Surprisingly, the competition effect in this case appears to be expressed primarily on the formation of new axon branches, rather than on branch stability.

Exogenous expression of the human inward rectifier K⁺ channel Kir2.1 has been shown to be a useful tool for electrically silencing individual neurons¹⁶. We used calcium imaging to test for the ability of Kir2.1 expression to suppress the excitability of zebrafish neurons. Figure 1a exemplifies the pattern of spontaneous and synchronized calcium spiking associated with spontaneous electrical activity in spinal cord neurons from larval zebrafish¹⁷. Figure 1b shows the effects of Kir2.1 expression on this calcium spiking. Whereas 41% of ventral lateral spinal neurons spiked in DsRed-expressing controls, only 6% ($P < 0.01$, binominal distribution statistics) spiked in those co-transfected with DsRed and Kir2.1. In animals expressing a non-conducting mutant version of Kir2.1 (Kirm)¹⁶, the percentage of active neurons was 50%, which was comparable to controls. These results indicate that Kir2.1 channel activity is effective in suppressing the electrical excitability of individual zebrafish neurons. Although optical imaging limitations precluded a direct test for silencing in transfected RGCs, it appears unlikely that effects of Kir2.1 in RGCs would be different from those observed in spinal cord neurons.

To express Kir2.1 in single RGCs, we isolated a 6-kilobase upstream sequence of the *brn3c* (*pou4f3*) gene (T. Xiao, T. Roeser, W. Staub and H. B., manuscript in preparation) as promoter. We chose zebrafish neuronal vesicle-associated membrane protein (VAMP) fused to green fluorescent protein (GFP) as an axon marker. Blastomere injection of *brn3c* promoter:GAL4 and a double UAS (the DNA recognition sequence that GAL4 binds to) cassette expressing Kir2.1 and VAMP-GFP led to just below 90% coexpression efficiency (Supplementary Fig. 1), which is consistent with an earlier report¹⁸. Most transfected animals expressed VAMP-GFP in a single RGC axon in one retina. We examined the effect of Kir2.1 expression on RGC axon arbor growth by measuring arbor length (as summed axon branch lengths) and complexity (as the number of branches per arbor) by two-photon microscopy *in vivo*. At 3 days after fertilization, when axons first start arborizing in the tectum, no differences were detected between

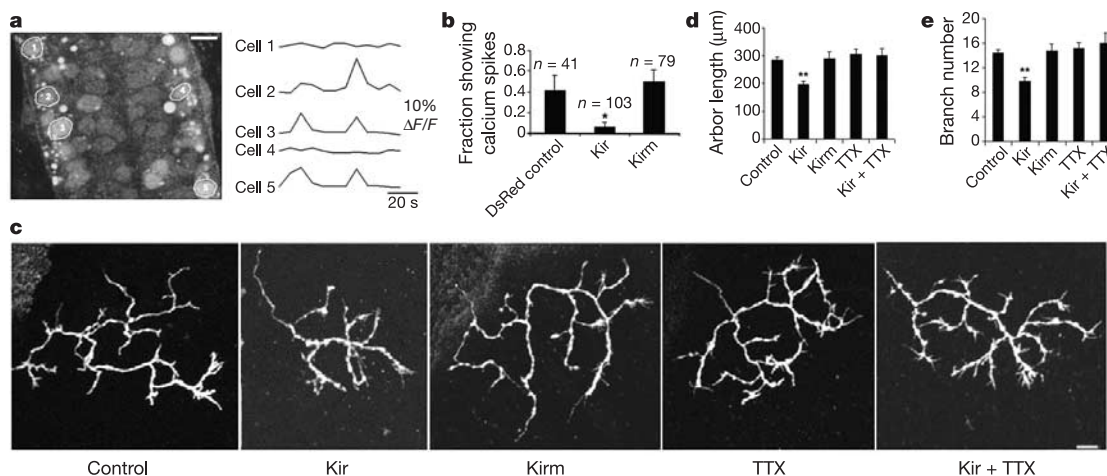


Figure 1 Kir2.1 overexpression inhibits zebrafish neuronal calcium spiking; expression in single RGCs inhibits axon arbor growth. **a**, Zebrafish ventral lateral spinal cord neurons exhibit spontaneous calcium spiking. The traces in the right panel show fluorescent intensity signals for the neurons marked by circles. The asterisk marks fluorescent cell debris. **b**, The fractions of neurons exhibiting calcium spiking ($n > 10$ animals for each

group). **c–e**, Sample axon arbor images (**c**), average axon arbor length (**d**) and branch number (**e**) for single transfected RGCs from animals at 5 days after fertilization. Control, $n = 34$; Kir, $n = 29$; Kirm, $n = 16$; TTX, $n = 19$; Kir + TTX, $n = 13$. Results are presented as means $\pm$ s.e.m.; asterisk indicates $P < 0.01$, double asterisk indicates $P < 0.001$. Scale bar, 5 μm .

Kir2.1-expressing (Kir) neurons and control neurons expressing only VAMP-GFP (data not shown). At 5 days after fertilization, however, Kir2.1 expression in single neurons caused a significant decrease in both arbor length (to 68% of control, $P < 0.001$) and arbor complexity (to 68% of control, $P < 0.001$) (Fig. 1c–e). In comparison, expression of the non-conducting Kir2.1 mutant (Kirm) had essentially no effect on either arbor length or complexity (Fig. 1c–e). It therefore seems that arbor growth inhibition by Kir2.1 is a result of Kir2.1 channel activity, acting through the suppression of electrical excitability.

The inhibition of axon growth by Kir2.1 channel activity could result directly from the suppression of RGC depolarization, or indirectly from blocking neurotransmitter release that is normally triggered by electrical excitation. To distinguish between these possibilities, we constructed a dominant-negative zebrafish VAMP-GFP protein (VAMPm) based on a mammalian dominant-negative VAMP mutant described previously¹⁹ (see Methods). Figure 2a, b shows results confirming the ability of VAMPm to block synaptic vesicle function with the use of the synaptic vesicle uptake marker FM1-43 (ref. 20). VAMPm overexpression suppresses vesicular FM1-43 uptake, reducing the fraction of

synaptic boutons exhibiting synaptic vesicle recycling in zebrafish motor neurons to 28% of control levels ($P < 0.001$).

VAMPm expression in single RGC axons did not affect arbor morphology at 3 days after fertilization (data not shown), but reduced axon arbor length and branch number to about 60% of control by 5 days after fertilization ($P < 0.001$, Fig. 2c–e). This growth inhibition is unlikely to be caused by VAMPm interference with SNARE proteins involved in intracellular trafficking, because SNARE interactions are pathway-specific *in vivo*²¹. Furthermore, VAMPm did not inhibit axon growth in conditions under which neural activity is globally blocked (see below). Because the inhibition of arbor growth by VAMPm resembles that caused by Kir2.1, most or all of the effects of electrical excitation might be due to depolarization-triggered neurotransmitter release.

To test for an activity-based competition rule, we measured the arbor growth of activity-suppressed cells when the activity of nearby cells was also suppressed in two different ways. We first performed imaging experiments in animals treated with an intraocular injection of tetrodotoxin (TTX) to block all RGC action potential firing. This TTX treatment did not affect the morphology of axons expressing VAMP-GFP, which is consistent with previous work¹⁴, but restored the growth of both Kir2.1-expressing (Fig. 1c–e) and VAMPm-expressing (Fig. 2c–e) RGC axon arbors to levels comparable ($P > 0.5$, that is, not significant) to the control. This

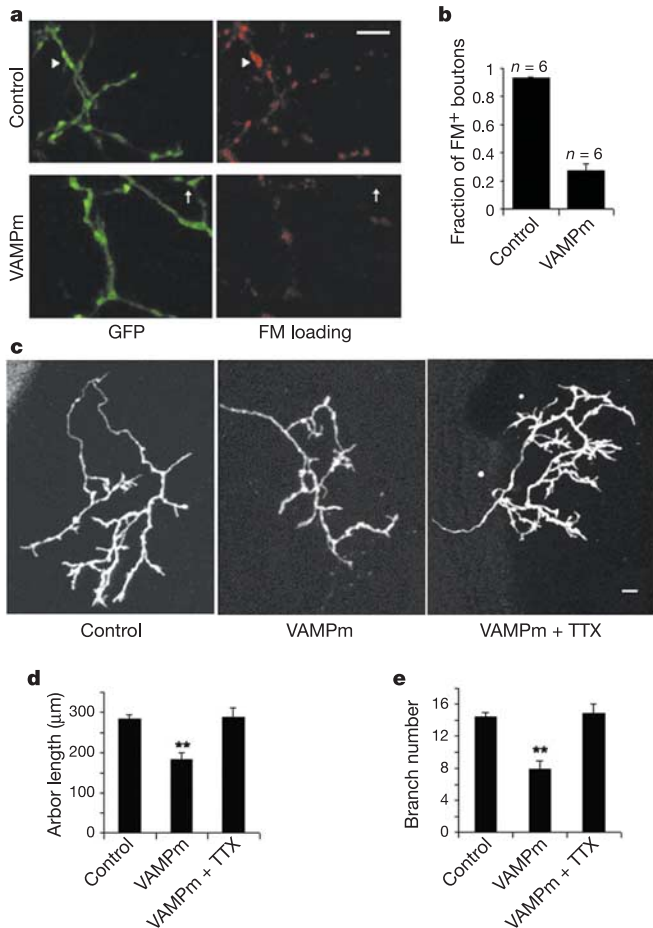


Figure 2 VAMPm overexpression suppresses presynaptic vesicle function in zebrafish neurons; expression in single RGCs inhibits axon arbor growth. **a, b**, Sample images (**a**) and quantification (**b**) of FM1-43 loading efficiency in VAMP-GFP (control) and VAMPm-expressing zebrafish spinal cord motor neurons. Arrowheads mark a FM1-43-loaded presynaptic bouton. Arrows mark a presynaptic bouton showing no FM1-43 uptake. $n = 6$ cells for each group. **c–e**, Sample axon arbor images (**c**), average axon arbor length (**d**) and branch number (**e**) for single transfected RGCs from animals at 5 days after fertilization. $n = 15$ for VAMPm, $n = 14$ for VAMPm + TTX. Results are presented as means $\pm$ s.e.m.; double asterisk indicates $P < 0.001$. Scale bar, 5 μm.

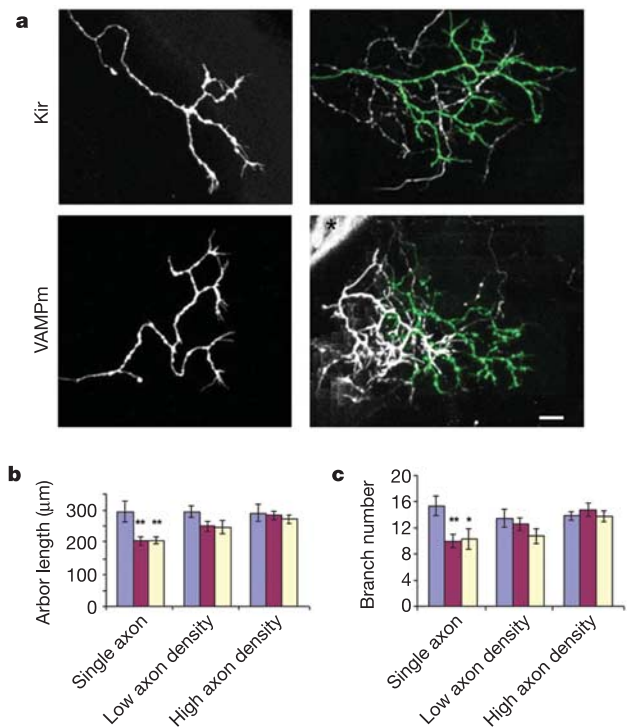


Figure 3 Transfection of multiple neighbouring RGC axons with Kir2.1 or VAMPm mitigates their growth-suppressing effects compared with transfection of a single axon. **a**, Sample images of electroporated RGCs at 5 days after fertilization. Single transfected axons are shown at the left, multiple transfected axons at the right. In each of the right panels one transfected axon arbor is highlighted in green. Skin fluorescence is marked by an asterisk. **b, c**, The axon arbor length (**b**) and branch number (**c**) for electroporated RGCs 5 days after fertilization. Axon densities between 0 and $0.095 \mu\text{m}^{-1}$ were classified as low, those between 0.095 and $0.19 \mu\text{m}^{-1}$ as high. Axon density was 0.089 for the highlighted Kir axon and 0.11 for the highlighted VAMPm axon in **a**. At least nine arbors were analysed for each condition. Colour coding in **b** and **c**: blue bars, control; maroon bars, Kir; yellow bars, VAMPm. Results are presented as means $\pm$ s.e.m.; asterisk indicates $P < 0.01$, double asterisk indicates $P < 0.001$. Scale bar, 5 μm.

pharmacological 'rescue' from the growth-inhibiting effects of single-cell activity inhibition indicates that competition from adjacent RGC axons, eliminated by the TTX treatment, might be responsible for the inhibition seen in the absence of TTX. However, TTX might diffuse from the eye into the brain and suppress activity in postsynaptic tectal cells, so it remains possible that an effect of direct competition between axons is distorted or obscured by silencing tectal cells with TTX.

In a second approach, which permitted a more selective suppression of activity in RGCs, we transfected groups of neighbouring RGCs by retinal electroporation. Sample images are shown in Fig. 3a. As expected if arbor growth is regulated by competition between neighbouring RGC axons, we observed a positive correlation between axon growth and the local density of axons expressing activity-suppressing constructs. To quantify the degree of overlap between transfected axons, each axon was assigned an 'axon density' value, based on the length of neighbouring transfected axon branches in its projection area. Figure 3b, c shows that growth inhibition was maximal in single transfected axons and returned almost to control levels at the highest axon densities achieved. The absence of an axon density effect for the control VAMP-GFP axons makes artificial effects of electroporation efficiency on axon growth unlikely. Comparisons with electron microscopic sections of tectal neuropil (S.J.S., unpublished data) and the stable *brn3c:mGFP* transgenic line (T. Xiao, T. Roeser, W. Staub and H.B., manuscript in preparation) indicate that even our highest densities of transfected axons might represent only a small fraction

of the axons converging into our imaging volumes. The axon density effect seen here indicates that the RGC subsets transfected by the retinal electroporation method might converge functionally in the tectum in some concerted way, predisposing them to strong growth competition.

Axon arbor growth is known to be a highly dynamic process: net arbor growth results from a balance between branch formation and elimination, with only a small fraction of new branches being maintained in the mature arbor²². To tease apart specific effects on arbor growth dynamics that underlie the activity-based competition, we collected time-lapse movies of single transfected RGC axons. Because the effects of competition on axon morphology are evident by 5 days after fertilization, we performed time-lapse imaging at 4–5 days after fertilization (100–112 h after fertilization).

We collected images at 2-min intervals for 40–50 min to capture rapid axon dynamics (Fig. 4a and Supplementary Movies 1–3). Expression of VAMPm and Kir2.1 in single RGCs substantially suppressed arbor motility (scored by summing the absolute values of length changes across all branches and time points) (Fig. 4b; Kir2.1 39% of control, VAMPm 59% of control, $P < 0.001$), and branch formation rates (Fig. 4c; Kir2.1 45% of control, VAMPm 56% of control, $P < 0.001$). About 95% of the newly formed branches were present in more than one image in the time-lapse sequences; the reduced branch formation rates observed were therefore unlikely to have been due to shortening of transient branch lifetimes to below detection.

In contrast, Kir2.1 and VAMPm expression in single RGCs did

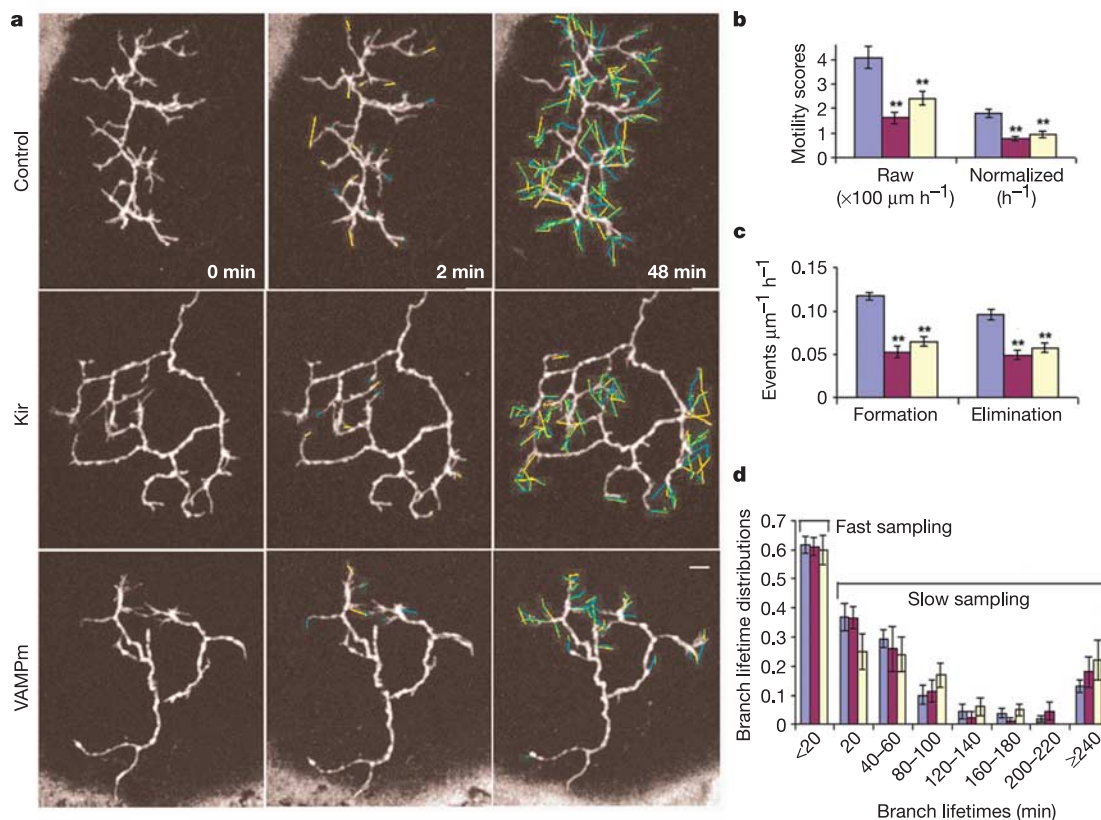


Figure 4 Selective activity suppression inhibits axon motility and branch formation rate with little effect on branch stability. **a**, Sample images taken from 2-min interval (fast-sampling) movies. The branch end positions between consecutive time points were connected by lines colour-coded for length change: yellow for increase, cyan for decrease. Traces in which branch extension and retraction overlap appear green. The middle panels show branch end position changes in the first 2 min; the right panels show the cumulative changes in 48 min. **b**, Motility scores from fast-sampling movies before

and after normalization to initial axon arbor length. $n = 9$ for control; $n = 8$ for Kir2.1 and VAMPm. **c**, Branch formation rates and elimination rates determined from fast-sampling movies. **d**, Lifetime distribution of newly formed branches for fast-sampling and slow-sampling (20-min interval) movies. At least 104 branches were analysed for each group. Colour coding in **b–d**: blue bars, control; maroon bars, Kir; yellow bars, VAMPm. Results are presented as means $\pm$ s.e.m.; double asterisk indicates $P < 0.001$. Scale bar, 5 μ m.

not affect axon branch stability. Of the new branches that formed at least 20 min before the end of imaging, similar proportions were eliminated within 20 min in control ($62 \pm 3\%$), Kir2.1 ($61 \pm 3\%$) and VAMPm ($60 \pm 5\%$) neurons. We also imaged axons at 20-min intervals for 8 h (Supplementary Movies 4–6) to determine branch stability over a longer time course. The lifetime distribution of branches formed during the first 4 h of imaging showed no statistical difference between control, Kir and VAMPm neurons (Fig. 4d), confirming a lack of activity-suppression effect on branch stability. Although Kir2.1 and VAMPm expression produced large reductions in branch elimination rates (Fig. 4c), the lack of any obvious competition effect on branch lifetime indicates that the reduced elimination rates might have been simply a consequence of the reductions in branch formation rates (that is, fewer transient branches were eliminated because fewer new branches formed).

Most of the cellular mechanisms at the heart of the competition rule we described here remain unknown. However, our results do indicate that many or all of the competition effects might be mediated by VAMP-dependent synaptic vesicle function. It is possible that some other consequence of synaptic vesicle function, such as co-transmitter release or membrane turnover, is involved, but it seems most likely that the released neurotransmitter itself mediates the observed effect, binding to nearby axon arbors or else to postsynaptic tectal cells or glial cells that feed back onto the RGC arbors through retrograde signals. Several plausible candidates for such a retrograde signal have been proposed, including neurotrophins, nitric oxide and arachidonic acid^{13,23,24}.

Discerning simple developmental rules that govern the construction of complex neural circuits is a central aspiration of developmental neuroscience. Here we present evidence for a specific competition rule in which the growth of an axon arbor is inhibited when its activity is suppressed below that of active neighbours. Our high-resolution time-lapse data further indicate that under the conditions examined such competition might govern the formation of new axonal branches but not branch stability. Compared with competition rules governing branch stabilization alone, competitive regulation of branch formation might increase the efficiency with which a dynamic axon arbor could 'explore' large neuropil volumes for appropriate synaptic partners. □

Methods

Ethical approval

All experimental protocols were approved by the Stanford University Animal Care and Use Committee.

Vector production

The *brn3c*:GAL4 vector was based on PB-GAL4VP16 (a gift from S. E. Fraser). In zebrafish neuronal VAMP-GFP (a gift from J. D. Jontes), GFP was tagged to the carboxyl terminus of VAMP. VAMP-GFP recycles between synaptic vesicles and the cell membrane, so it distributes in both a punctate and a membrane-targeted pattern, strongly labelling axon branch tips. To generate VAMPm, mutations of VAMP E81A, S82A were introduced to VAMP-GFP by polymerase chain reaction. Human Kir2.1 and the non-conducting mutant, in which the pore-region amino-acid motif GYG was mutated to AAA, were gifts from V. N. Murthy and E. Marban. A Myc tag was added to the amino terminus of Kir2.1 by polymerase chain reaction. UAS:VAMP-GFP:pA was inserted before the UAS sequence of UAS:Kir2.1:pA so that RNA transcription proceeded in the same direction. pA refers to the polyadenylation signal from Clontech PEGFP-N1 as a *NotI*/*Afl*III fragment. The pA following VAMP-GFP was inverted during subcloning, inverting the two original AATAAA motifs and creating a new one. Double UAS constructs and *brn3c* promoter:GAL4 were linearized for injection.

Embryo maintenance

Embryos were raised in the dark at 28 °C in the presence of 150 μ M 1-phenyl-2-thiourea to prevent pigment formation. The RGC axon morphology of animals raised in the dark was comparable to that of animals raised on a 14 h:10 h light–dark cycle at 5 days after fertilization (axon arbor length $295 \pm 23 \mu$ m, branch number 13.7 ± 1.5 , $n = 16$). For TTX treatment, larvae received an intraocular injection of 8–10 nl of 0.6 mM TTX (Calbiochem) in Ringer solution. Injection at 2 or 3 days after fertilization led to similar effects on RGC axon development, and the data were pooled for analysis.

RGC axon imaging and data analysis

Zebrafish larvae were immobilized in agarose. Z-series stacks were collected at 1.5- μ m intervals. Because RGC axons arborizing in the anterior tectum tend to have small projection areas and short arbor lengths, only axons arborizing in the posterior two-thirds of the tectum were considered. Axons with a GFP intensity higher than double the auto-fluorescence of skin were excluded from analysis. VAMP-GFP expression at the levels used here did not affect axon morphology compared with axons expressing GFP (axon arbor length $2,875 \pm 24 \mu$ m, branch number 14.6 ± 0.9 , $n = 10$). For time-lapse analysis, axons transfected by blastomere injection with axon arbor lengths of 200–300 μ m were chosen. For image analysis, processes at least 2 μ m long were considered branches. Growth cones bearing multiple filopodia were considered single branches. Axons at 5 days after fertilization were reconstructed three-dimensionally. For time-lapse analysis, maximum-intensity projections were used. The Kolmogorov–Smirnov test was used for statistics unless otherwise noted.

Calcium green-1 imaging

Calcium green-1 dextran (10 mM; 10 kDa; Molecular Probes) was injected into embryos together with α -tubulin promoter:GAL4 and UAS:DsRed or DsRed, Kir2.1 double-UAS construct. Larvae at 24–27 h after fertilization were immobilized in 100 μ g ml^{−1} curare for imaging. Images were collected from the ventral spinal cord (within 20 μ m from the notochord) at 10-s intervals for 2 min. (Optical artefacts due to lens and pigment precluded effective calcium imaging from RGC cell bodies in the retina. No method tried permitted the imaging of useful calcium signals from RGC axons.) Images were analysed with Openview software (authored by N. Ziv, Technion, Haifa, Israel). Ventral lateral spinal cord neurons showing 10% or more increases in fluorescent intensity in cell bodies were considered to be spiking. Neurons that fired at least once during the imaging interval were scored as active.

FM1-43 loading assay

FM1-43 loading was performed in zebrafish larvae 3 days after fertilization as described in ref. 20, except that the FM1-43 preloading incubation was omitted and the incubation in Advasep-7 after loading was extended to 15 min. For data analysis, regions of the motor neuron axons with a GFP intensity higher than double the intensity of shaft region were considered presynaptic boutons. FM1-43 puncta with an intensity of 80–255 and a size of 0.5–2 μ m were considered synaptic.

Retinal electroporation

Electroporation was performed as described in ref. 25. DNA plasmids were injected into the intraocular space of larvae at 2 days after fertilization. Three or four square 15-ms current pulses of 0.3 mA were delivered at 1-s intervals between a fine wire electrode placed near the injected eye and a distant ground electrode. Arbor length and branch number were measured from three-dimensional images. To measure axon density, Z-series stacks consisted of minimal numbers of sections containing the axon arbors of interest. Neighbouring transfected axon branches contained within this Z stack and the projection area of the arbor of interest were traced. Their summed length divided by the projection area produced the axon density value for the arbor of interest.

Received 21 April 2004; accepted 28 January 2005; doi:10.1038/nature03409.

- Sanes, J. R. & Lichtman, J. W. Development of the vertebrate neuromuscular junction. *Annu. Rev. Neurosci.* **22**, 389–442 (1999).
- Katz, L. C. & Shatz, C. J. Synaptic activity and the construction of cortical circuits. *Science* **274**, 1133–1138 (1996).
- Zhang, L. I. & Poo, M. M. Electrical activity and development of neural circuits. *Nature Neurosci.* **4** (Suppl), 1207–1214 (2001).
- Buffelli, M. *et al.* Genetic evidence that relative synaptic efficacy biases the outcome of synaptic competition. *Nature* **424**, 430–434 (2003).
- Lichtman, J. W. & Balice-Gordon, R. J. Understanding synaptic competition in theory and in practice. *J. Neurobiol.* **21**, 99–106 (1990).
- Shatz, C. J. & Stryker, M. P. Ocular dominance in layer IV of the cat's visual cortex and the effects of monocular deprivation. *J. Physiol. (Lond.)* **281**, 267–283 (1978).
- Stryker, M. P. & Harris, W. A. Binocular impulse blockade prevents the formation of ocular dominance columns in cat visual cortex. *J. Neurosci.* **6**, 2117–2133 (1986).
- Antonini, A. & Stryker, M. P. Rapid remodeling of axonal arbors in the visual cortex. *Science* **260**, 1819–1821 (1993).
- Yu, C. R. *et al.* Spontaneous neural activity is required for the establishment and maintenance of the olfactory sensory map. *Neuron* **42**, 553–566 (2004).
- Zhao, H. & Reed, R. R. X inactivation of the OCN1 channel gene reveals a role for activity-dependent competition in the olfactory system. *Cell* **104**, 651–660 (2001).
- Ruthazer, E. S., Skerfving, C. J. & Cline, H. T. Control of axon branch dynamics by correlated activity *in vivo*. *Science* **301**, 666–70 (2003).
- Schmidt, J. T., Buzzard, M., Borress, R. & Dhillon, S. MK801 increases retinotectal arbor size in developing zebrafish without affecting kinetics of branch elimination and addition. *J. Neurobiol.* **42**, 303–314 (2000).
- Schmidt, J. T., Fleming, M. R. & Leu, B. Presynaptic protein kinase C controls maturation and branch dynamics of developing retinotectal arbors: possible role in activity-driven sharpening. *J. Neurobiol.* **58**, 328–340 (2004).
- Gnueghe, L., Schmid, S. & Neuhaus, S. C. Analysis of the activity-deprived zebrafish mutant macho reveals an essential requirement of neuronal activity for the development of a fine-grained visuotopic map. *J. Neurosci.* **21**, 3542–3548 (2001).
- Johnson, F. A., Dawson, A. J. & Meyer, R. L. Activity-dependent refinement in the goldfish retinotectal system is mediated by the dynamic regulation of processes withdrawal: an *in vivo* imaging study. *J. Comp. Neurol.* **406**, 548–562 (1999).
- Burrone, J., O'Byrne, M. & Murthy, V. N. Multiple forms of synaptic plasticity triggered by selective suppression of activity in individual neurons. *Nature* **420**, 414–418 (2002).

17. Drapeau, P. *et al.* Development of the locomotor network in zebrafish. *Prog. Neurobiol.* **68**, 85–111 (2002).
18. Koster, R. W. & Fraser, S. E. Tracing transgene expression in living zebrafish embryos. *Dev. Biol.* **233**, 329–346 (2001).
19. Sorensen, J. B. *et al.* The SNARE protein SNAP-25 is linked to fast calcium triggering of exocytosis. *Proc. Natl Acad. Sci. USA* **99**, 1627–1632 (2002).
20. Li, W., Ono, F. & Brehm, P. Optical measurements of presynaptic release in mutant zebrafish lacking postsynaptic receptors. *J. Neurosci.* **23**, 10467–10474 (2003).
21. Scales, S. J. *et al.* SNAREs contribute to the specificity of membrane fusion. *Neuron* **26**, 457–464 (2000).
22. Hua, J. Y. & Smith, S. J. Neural activity and the dynamics of central nervous system development. *Nature Neurosci.* **7**, 327–332 (2004).
23. Cogen, J. & Cohen-Cory, S. Nitric oxide modulates retinal ganglion cell axon arbor remodeling *in vivo*. *J. Neurobiol.* **45**, 120–133 (2000).
24. Cohen-Cory, S. & Fraser, S. E. Effects of brain-derived neurotrophic factor on optic axon branching and remodelling *in vivo*. *Nature* **378**, 192–196 (1995).
25. Haas, K., Jensen, K., Sin, W. C., Foa, L. & Cline, H. T. Targeted electroporation in *Xenopus* tadpoles *in vivo*—from single cells to the entire brain. *Differentiation* **70**, 148–154 (2002).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank T. R. Clandinin, R. W. Tsien, R. W. Aldrich, L. Luo and the Smith laboratory for comments on the manuscript, and C. M. Niell for developing the Matlab routines used in image analysis. We thank T. Roeser for isolating the *brn3c* promoter. The US National Institutes of Health and the Vincent Coates Foundation provided financial support. J.Y.H. was supported by a Stanford Graduate Fellowship and a Coates Foundation Fellowship. M.C.S. was supported by a predoctoral fellowship from the American Heart Association.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.Y.H. (huay@stanford.edu).

Hypothalamic K_{ATP} channels control hepatic glucose production

Alessandro Pocai¹, Tony K. T. Lam¹, Roger Gutierrez-Juarez¹, Silvana Obici¹, Gary J. Schwartz¹, Joseph Bryan², Lydia Aguilar-Bryan² & Luciano Rossetti¹

¹Departments of Medicine, Neuroscience and Molecular Pharmacology, Diabetes Research Center, Albert Einstein College of Medicine, Bronx, New York 10461, USA

²Department of Medicine, Baylor College of Medicine, Houston, Texas 77030, USA

Obesity is the driving force behind the worldwide increase in the prevalence of type 2 diabetes mellitus^{1,2}. Hyperglycaemia is a hallmark of diabetes and is largely due to increased hepatic gluconeogenesis³. The medial hypothalamus is a major integrator of nutritional and hormonal signals^{1,2,4}, which play pivotal roles not only in the regulation of energy balance but also in the modulation of liver glucose output^{5,6}. Bidirectional changes in hypothalamic insulin signalling therefore result in parallel changes in both energy balance^{7–10} and glucose metabolism⁵. Here we show that activation of ATP-sensitive potassium (K_{ATP}) channels¹¹ in the mediobasal hypothalamus is sufficient to lower blood glucose levels through inhibition of hepatic gluconeogenesis. Finally, the infusion of a K_{ATP} blocker within the mediobasal hypothalamus, or the surgical resection of the hepatic branch of the vagus nerve, negates the effects of central insulin and halves the effects of systemic insulin on hepatic glucose production. Consistent with these results, mice lacking the SUR1 subunit of the K_{ATP} channel¹² are resistant to the inhibitory action of insulin on gluconeogenesis. These findings suggest that activation of hypothalamic K_{ATP} channels normally restrains hepatic gluconeogenesis, and that any alteration

within this central nervous system/liver circuit can contribute to diabetic hyperglycaemia.

ATP-sensitive potassium (K_{ATP}) channels are expressed in the hypothalamus¹³ and can be activated by insulin (and leptin) in selective hypothalamic neurons^{14,15}. However, the functional role of insulin activation of hypothalamic K_{ATP} channels remains obscure. To investigate whether direct activation of central K_{ATP} channels is sufficient to reproduce the potent effects of insulin on blood glucose levels, on glucose production, and on hepatic gluconeogenesis, we infused the K_{ATP} channel activator diazoxide in the third cerebral ventricle (ICV) of conscious rats. Central administration of diazoxide lowered blood glucose levels (Fig. 1a). To examine the mechanisms by which central activation of K_{ATP} channels decreases blood glucose, we combined ICV infusions of vehicle or diazoxide with systemic pancreatic insulin-clamp studies (Supplementary Fig. S1a). In the presence of basal circulating insulin levels (~20 µU ml⁻¹), glucose infusion was required to prevent hypoglycaemia following central administration of diazoxide (Fig. 1b). We next assessed glucose kinetics to establish whether the increased requirement for glucose infusion in response to central activation of K_{ATP} channels is due to stimulation of glucose uptake or to inhibition of glucose production. ICV diazoxide markedly and significantly decreased glucose production (by 45 ± 4%; Fig. 1b), but the rate of glucose uptake was not significantly affected (Supplementary Fig. S1b). Thus, central stimulation of K_{ATP} channels lowers blood glucose by inhibiting glucose production.

Glucose production represents the net increase in glucosyl units derived from gluconeogenesis and glycogenolysis. However, a portion of glucose entering the liver as a result of glucose phosphorylation is also a substrate for dephosphorylation by glucose-6-phosphatase (G6Pase), creating a futile cycle of glucose cycling (Supplementary Fig. S1a). In order to further delineate the mechanisms by which central activation of K_{ATP} channels modulates glucose homeostasis, we estimated the *in vivo* flux through G6Pase and the relative contribution of gluconeogenesis and glycogenolysis to glucose output. ICV diazoxide infusion decreased the flux through G6Pase (Supplementary Fig. S1c) in parallel with its effects on glucose production (Fig. 1b). Importantly, the decrease in glucose production was largely accounted for by marked inhibition of gluconeogenesis (Fig. 1b), while the rate of glycogenolysis was not significantly decreased (Supplementary Fig. S1d). Real-time polymerase chain reaction (PCR) analysis showed that ICV diazoxide infusion markedly decreased messenger RNA levels of liver *G6Pase* (also known as *G6pc* in mouse and rat) and of the gluconeogenic gene *Pepck* (*phosphoenolpyruvate kinase*, also known as *Pck1* in mouse and rat) (Fig. 1c). Thus, direct activation of central K_{ATP} channels was sufficient to mimic the actions of insulin on gluconeogenesis, the *in vivo* fluxes through G6Pase, and the hepatic expression of *G6Pase* and *Pepck*.

The potent metabolic effects of diazoxide could be mediated by its activation of K_{ATP} channels anywhere within the central nervous system (CNS)¹⁶. To examine the anatomical localization of these effects we infused a 15-fold lower dose of diazoxide (Fig. 1e) bilaterally within the parenchyma of the medial hypothalamus (Supplementary Fig. S2a). Intrahypothalamic infusion of diazoxide lowered blood glucose levels (Fig. 1d). This hypoglycaemic effect was due to marked suppression of glucose production (Fig. 1d). To examine whether the effects of ICV insulin on liver glucose homeostasis⁵ are also centred in an overlapping hypothalamic area, we next performed intrahypothalamic infusions of insulin (Fig. 1e) at a dose 15-fold lower than that previously used in ICV experiments⁵. Intrahypothalamic insulin reproduced the potent effects of ICV insulin on blood glucose concentration and on glucose infusion and gluconeogenesis (Fig. 1e). Thus, activation of either K_{ATP} channels or of insulin signalling within the medial hypothalamus is sufficient to decrease blood glucose levels through

suppression of glucose production, of hepatic gluconeogenesis, and of *Pepck* and *G6Pase* expression. These 'gain-of-function' experiments suggest that modulation of K_{ATP} channel activity within the medial hypothalamus can have a large effect on liver glucose homeostasis.

Sulphonylureas (K_{ATP} channel blockers) abolish the activation of hypothalamic K_{ATP} channels by insulin and leptin^{14,15}. Paired groups of rats received ICV infusions of vehicle, insulin, insulin and K_{ATP} blocker, or K_{ATP} blocker alone (Fig. 1f). In the presence of near-basal circulating insulin levels, glucose infusion was required

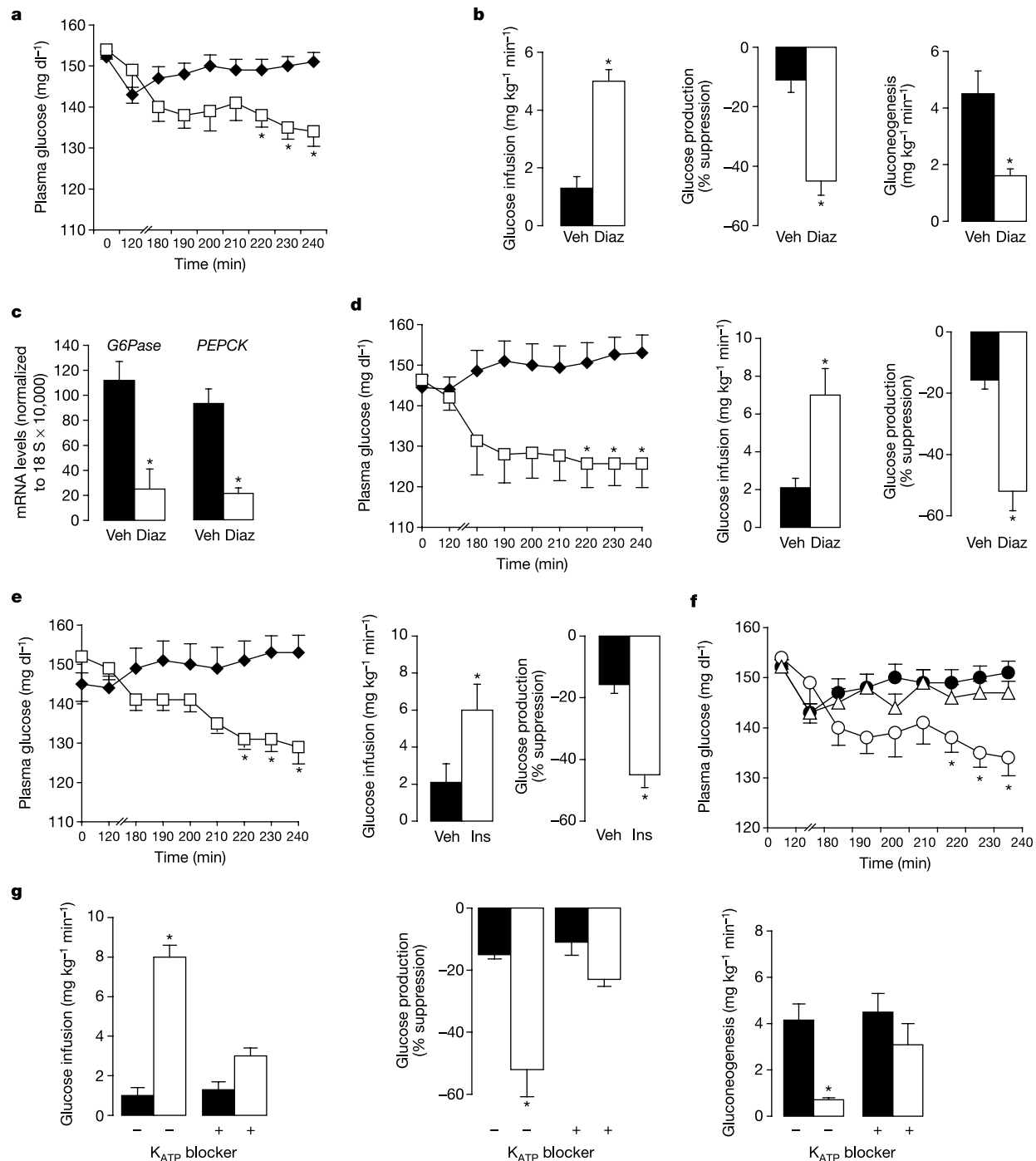


Figure 1 Activation of hypothalamic K_{ATP} channels lowers plasma glucose by inhibiting glucose production. **a**, Glucose levels during ICV diazoxide ($n = 6$; open squares) or vehicle ($n = 4$; filled diamonds) administration. **b**, Glucose infusion, glucose production and gluconeogenesis during pancreatic insulin clamps with ICV diazoxide (Diaz) or vehicle (Veh). **c**, *G6Pase* and *Pepck* mRNA. Gene copy number is normalized to the copy number of the housekeeping 18S gene ($\times 10,000$). **d**, Glucose levels during intrahypothalamic (IH) diazoxide ($n = 4$; open squares) and vehicle ($n = 5$; filled diamonds) administration, and glucose infusion and production during pancreatic clamps. **e**, Glucose levels during IH

insulin ($n = 4$; open squares) and vehicle ($n = 4$; filled diamonds) administration, and glucose infusion and production during pancreatic clamps. **f**, ICV K_{ATP} blocker ($n = 5$; open triangles) negates the glucose-lowering effect of ICV insulin ($n = 7$; open circles) compared with ICV vehicle ($n = 5$; filled circles) and K_{ATP} blocker alone ($n = 5$; not shown). **g**, Effect of ICV insulin (open bars) or vehicle (filled bars) on glucose infusion, glucose production and gluconeogenesis, in the presence and absence of a K_{ATP} blocker. Asterisk denotes $P < 0.05$ versus vehicle or versus insulin + K_{ATP} blocker. All values are mean $\pm$ s.e.m.

to prevent hypoglycaemia after central administration of insulin (Fig. 1g). ICV infusion of the K_{ATP} blocker glibenclamide alone did not modify the rate of glucose infusion compared with ICV vehicle (Fig. 1g). Importantly, central administration of insulin did not alter glucose requirements when the K_{ATP} blocker was co-infused. In the presence of basal insulin levels ($\sim 20 \mu\text{U ml}^{-1}$), ICV insulin markedly and significantly decreased glucose production (by $52 \pm 8\%$, Fig. 1g) and the flux through G6Pase (Supplementary Fig. S1f). This decrease was negated by the ICV co-infusion of K_{ATP} blocker and was largely accounted for by marked inhibition of gluconeogenesis (Fig. 1g), as the rate of glycogenolysis was not significantly decreased (Supplementary Fig. S1h). Importantly, in the presence of ICV glibenclamide, central insulin did not affect hepatic glucose fluxes. Central stimulation of insulin action also resulted in a marked K_{ATP} channel-dependent decrease in liver *G6Pase* and *Pepck* mRNA levels (Supplementary Fig. S1i). Thus, the ICV infusion of a K_{ATP} blocker completely prevented the central effects of insulin on liver glucose homeostasis.

Genetic and electrophysiological evidence have indicated the presence of K_{ATP} channels in selective hypothalamic neurons. These channels have an octameric structure similar to that of peripheral K_{ATP} channels, with a K^+ inward rectifier subunit ($K_{IR6.1}$ or 6.2), and a sulphonylurea receptor, SUR1 or SUR2 (refs 11, 17). To identify the composition of K_{ATP} channels within the hypothalamus, we investigated their expression in key hypothalamic nuclei. Messenger RNA for both *Sur1* and *Sur2* (also known as *Abcc8* and *Abcc9*, respectively) is detectable in the medial hypothalamus (arcuate nuclei) (Supplementary Fig. S2b). Hypothalamic neurons expressing K_{ATP} channels are targets of insulin¹⁵, and their high sensitivity to diazoxide and low

concentrations of sulphonylurea¹⁸ suggest that SUR1 is a component of these insulin-responsive K_{ATP} channels¹¹.

The absence of SUR1-containing K_{ATP} channels has been shown to lead to an increase in membrane potential that cannot be suppressed by diazoxide¹². As our findings in rats suggest that the neuronal hyperpolarization induced by hypothalamic administration of insulin or diazoxide results in suppression of gluconeogenesis, we next investigated whether the ability of insulin to restrain hepatic gluconeogenesis is selectively impaired in SUR1 null (*Sur1*^{-/-}) mice. To this end, we performed insulin-clamp studies in conscious *Sur1*^{-/-} and wild-type mice (Fig. 2a). *Sur1*^{-/-} mice showed hepatic (Fig. 2b) but not peripheral (Supplementary Fig. S2c) insulin resistance, with an approximately twofold increase in liver glucose production compared with wild-type mice. This increase in glucose production was largely due to a marked increase in gluconeogenesis, while glycogenolysis was not significantly altered (Fig. 2b). Thus, we propose that insulin activation of SUR1-containing K_{ATP} channels within the hypothalamus is required to restrain hepatic gluconeogenesis. This is the first demonstration of a defect in hepatic insulin action in SUR1 null mice, and it stands in contrast with the increased insulin sensitivity reported in $K_{IR6.2}$ and SUR2 null mice^{19,20}. Together with the results from 'gain-of-function' experiments, these pharmacological and genetic 'loss-of-function' experiments in rats and mice (respectively) indicate that SUR1-containing K_{ATP} channels within the medial hypothalamus have an important role in the regulation of liver glucose homeostasis.

Hypothalamic centres participate in the short-term regulation of ingestive behaviour by means of descending neural connections to the caudal brainstem, leading to activation of vagal input to the

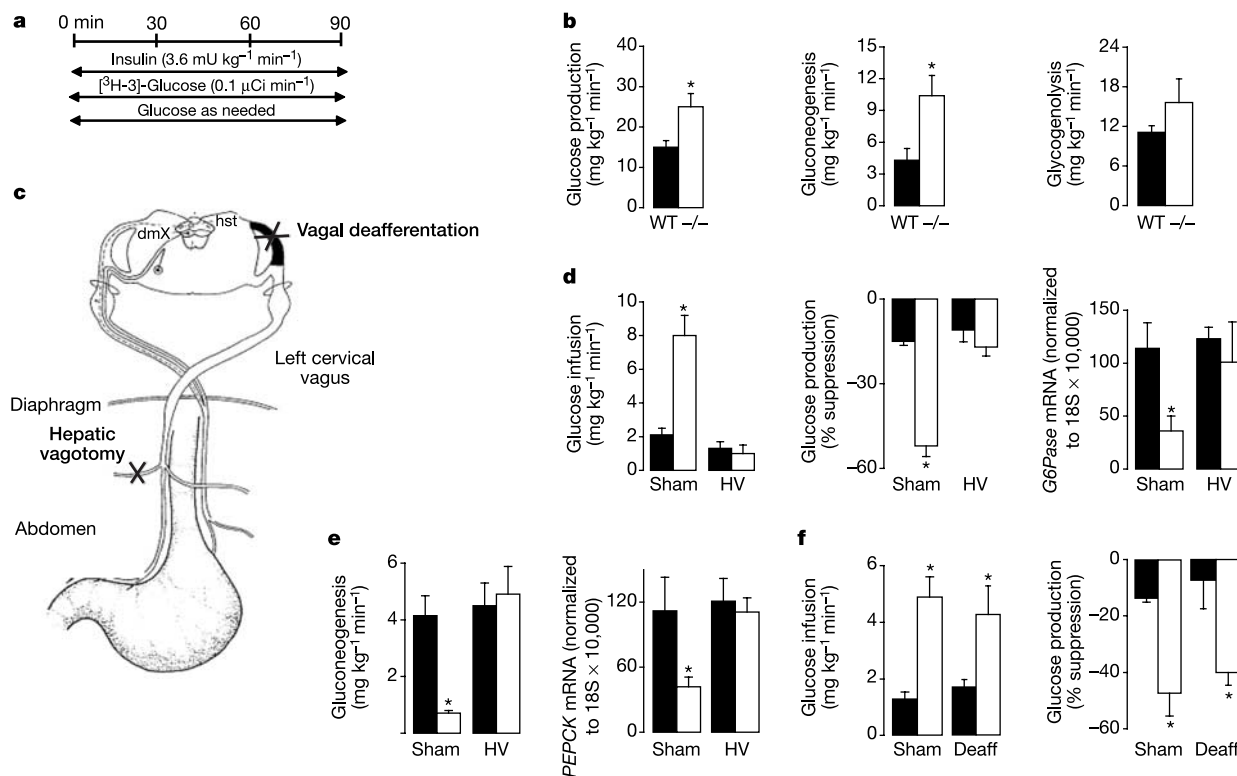


Figure 2 Role of SUR1 and efferent vagus in hepatic insulin action. **a**, Schematic of pancreatic clamp procedure in *Sur1*^{-/-} mice. **b**, Glucose production, gluconeogenesis and glycogenolysis in wild-type (WT, $n = 8$) and *Sur1*^{-/-} ($n = 6$) mice. **c**, Hepatic vagotomy and vagal deafferentation. **d**, ICV insulin (open bars) but not vehicle (filled bars) increases glucose infusion and inhibits glucose production and *G6Pase* expression in

sham-operated ($n = 11$; Sham) but not in vagotomized ($n = 11$; HV) rats. **e**, HV negates the effect of insulin on gluconeogenesis and *Pepck* mRNA. **f**, Vagal deafferentation (Deaff; $n = 8$) failed to alter the effects of ICV insulin (open bars) on glucose infusion and glucose production. Asterisk denotes $P < 0.05$ versus vehicle controls. All values are mean $\pm$ s.e.m.

gastrointestinal tract^{4,16}. Because autonomic neural input to the liver can also rapidly modulate liver metabolism²¹, we next asked whether central administration of insulin decreases glucose production and the expression of *G6Pase* and *Pepck* through activation of hepatic efferent vagal fibres. We tested the effects of central administration of insulin in rats with selective hepatic branch vagotomy (HV) or a sham operation (sham) (Fig. 2c–e). In the presence of approximately basal circulating insulin levels, glucose infusion was required to prevent hypoglycaemia following ICV insulin in the sham but not in the HV rats (Fig. 2d). ICV insulin significantly decreased glucose production (by $48 \pm 3\%$) in sham but not in HV rats (Fig. 2d). Glucose uptake was not affected (Supplementary Fig. S2e). ICV insulin markedly decreased the flux through *G6Pase* in the sham rats (Supplementary Fig. S2f). The decrease in glucose production was largely accounted for by inhibition of gluconeogenesis (Fig. 2e). Importantly, hepatic branch vagotomy negated the decreases in the *in vivo* flux of *G6Pase* and gluconeogenesis, and in the hepatic expression of *G6Pase* and *Pepck* in response to central administration of insulin (Fig. 2d, e).

The hepatic branch of the vagus nerve comprises both efferent and afferent fibres, and its resection abolished the hepatic effects of ICV insulin. To address the potential role of redundant hepatic branch vagal fibres in mediating the hypoglycaemic effects of ICV insulin, we performed additional experiments in animals with selective vagal deafferentation. These animals have intact descending efferent fibres to the liver, but all the vagal afferents supplying the hepatic vagal branch are resected at the site of entry to the brainstem (Fig. 2c). Vagal deafferentation did not alter the ability of central insulin to lower blood glucose levels and to suppress glucose production (Fig. 2f). Therefore, efferent vagal input to the liver is required for the inhibition of glucose production following central administration of insulin, but the afferent input from the hepatic branch of the vagus nerve to the brainstem is not required.

To estimate the contribution of the hepatic branch of the vagus to the overall effects of circulating insulin on liver glucose homeostasis, we generated physiological increases in circulating insulin levels in sham and HV rats using the pancreatic insulin-clamp technique (Fig. 3a). In the presence of similar ~3-fold increases

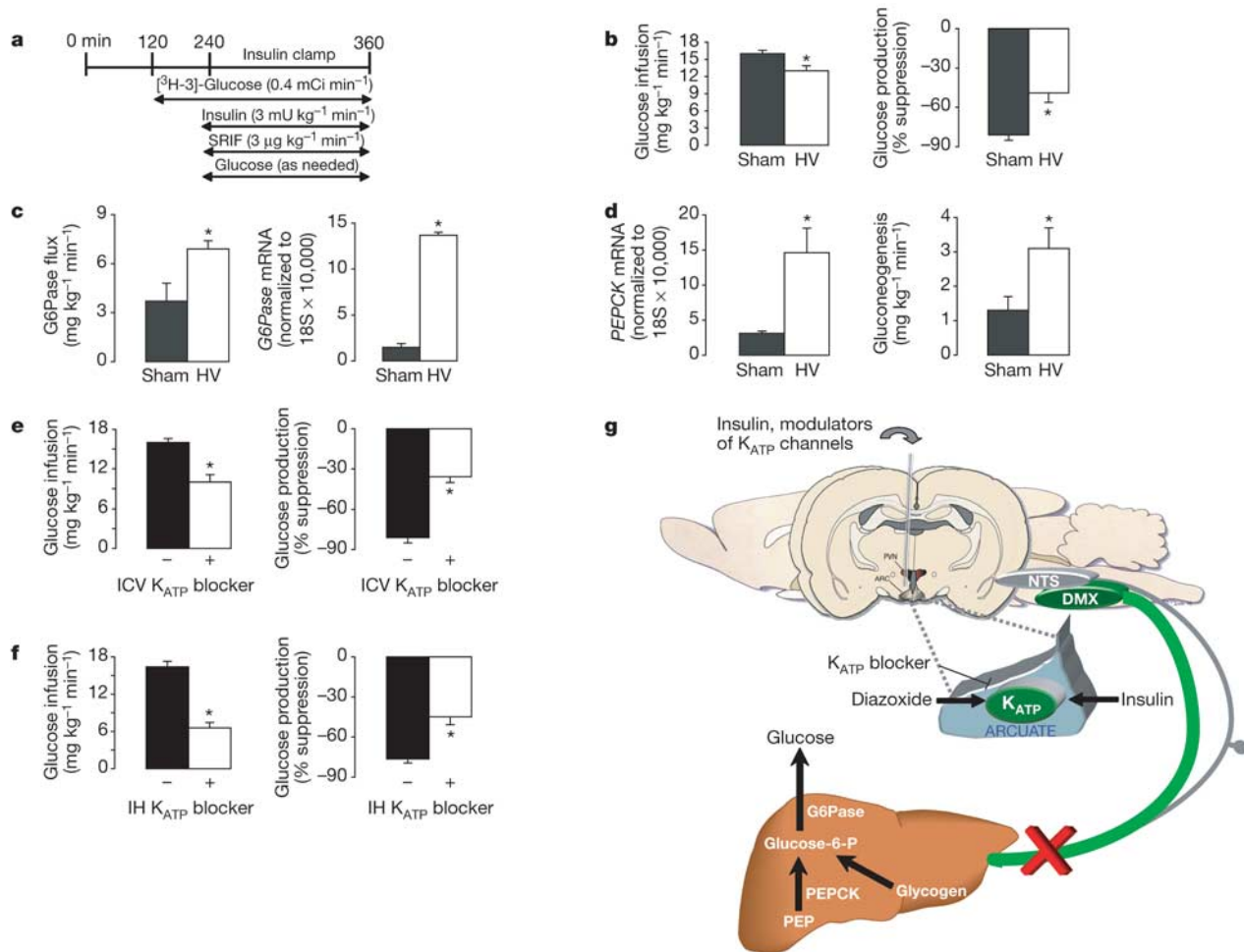


Figure 3 Hepatic vagus and activation of hypothalamic K_{ATP} channels are required for the effect of physiological hyperinsulinaemia on glucose production. **a**, Insulin-clamp procedure. SRIF, somatostatin. **b**, Glucose infusion and glucose production in rats with hepatic vagotomy (HV, $n = 13$; open bars), or a sham operation (Sham, $n = 13$; filled bars). **c**, d, HV impairs the effects of insulin on *G6Pase* flux and *G6Pase* mRNA (**c**) and on gluconeogenesis and *Pepck* mRNA (**d**). **e**, **f**, Administration of a K_{ATP} blocker ICV (**e**, $n = 4$) or IH (**f**, $n = 4$) impairs the effects of systemic insulin on glucose infusion and glucose

production. **g**, Longitudinal view of a rat brain. A coronal section of the brain at the level of the arcuate nuclei (foreground) shows the positioning of one of the infusion cannulae used for the IH studies. Activation of neuronal projections from the hypothalamus to brainstem nuclei (DMX, dorsal motor vagal nucleus; NTS, nucleus of the solitary tract) generates an efferent vagal impulse to the liver required for inhibition of glucose production. Asterisk denotes $P < 0.05$ versus sham. All values are mean $\pm$ s.e.m.

(Supplementary Table 4) in plasma insulin levels, higher rates of glucose infusion were required in sham compared with HV rats to prevent hypoglycaemia (Fig. 3b) because glucose production was inhibited more in sham rats ($80 \pm 8\%$) than in HV rats ($38 \pm 11\%$; Fig. 3b). We next estimated the *in vivo* flux through G6Pase and the relative contributions of gluconeogenesis and glycogenolysis (Supplementary Fig. S3d) to glucose output. The flux through G6Pase (Fig. 3c), gluconeogenesis (Fig. 3d), and the hepatic expression of *G6Pase* and *Pepck* were all increased in HV rats compared with sham controls (Fig. 3c, d). Hepatic branch vagotomy therefore interferes with the inhibitory effects of systemic insulin administration. Finally, the effects of physiological increases in plasma insulin levels on glucose infusion and glucose production were markedly and similarly impaired by ICV (Fig. 3e) or intrahypothalamic (Fig. 3f) infusion of a K_{ATP} blocker. Blocking the effects of insulin on hypothalamic K_{ATP} channels or on the hepatic branch of the vagus therefore results in loss of approximately half of the inhibitory action of circulating insulin on glucose production.

Together with the observation of hepatic insulin resistance in SUR1 null mice, our findings establish a physiological role for this insulin-driven brain–liver circuit (Fig. 3g). Physiological increases in circulating insulin levels activate a neuronal circuit that is required in order to restrain hepatic gluconeogenesis. Because increased gluconeogenesis is a main cause of fasting and postprandial hyperglycaemia in type II diabetes³, improving hypothalamic insulin signalling could become a new therapeutic strategy for this disease. There is increasing evidence to show that nutritional and hormonal signals act on the medial hypothalamus to curtail both exogenous and endogenous input of nutrients^{22–24}. The remarkable similarities between the brain–liver neural circuit activated in response to changes in hypothalamic lipid oxidation^{6,24,25} and in response to activation of hypothalamic insulin signalling suggest a convergence of these nutritional signals, perhaps at the level of hypothalamic K_{ATP} channels.

Bidirectional changes in the activity of hypothalamic K_{ATP} channels rapidly modulate circulating glucose levels by altering autonomic input to the liver. This new and important role for neuronal K_{ATP} channels in the regulation of glucose homeostasis adds to their well-established role in the regulation of pancreatic β -cell function¹⁷. K_{ATP} channels therefore play a pleiotropic role in the modulation of blood glucose levels—their activation in the hypothalamus decreases hepatic gluconeogenesis, and their activation in pancreatic β -cells decreases insulin secretion. The central and peripheral actions of K_{ATP} channels might be designed to balance each other in order to maintain glucose homeostasis. Any disruption in the equilibrium between these regulatory circuits is likely to result in altered glucose homeostasis. □

Methods

Animal preparation for *in vivo* experiments

We studied 89 10-week-old male Sprague–Dawley rats (Charles River Breeding Laboratories). Rats were housed in individual cages and subjected to a standard light–dark cycle. Three weeks before the *in vivo* studies, we used stereotaxic surgery to implant chronic catheters in the third cerebral ventricle²⁵ or intrahypothalamically²⁶. One week before pancreatic insulin-clamp experiments, rats received additional catheters in the right internal jugular vein and left carotid artery²⁵ (Supplementary Fig. S1a).

Adult male *Sur1*^{−/−} ($n = 8$; ref. 12) and wild-type ($n = 8$) littermate mice (27–32 g) were anaesthetized with chloral hydrate (400 mg kg^{−1} body weight intraperitoneally) and catheterized through the right internal jugular vein as previously described²⁷. The venous catheter was used for infusions, and blood samples were collected from the tail vein²⁷.

In vivo studies

For the studies in rats, the metabolic experiments were performed approximately 3 weeks after stereotaxic surgery, after complete recovery from the operation (Supplementary Fig. S1a). All solutions were diluted with artificial cerebral spinal fluid (aCSF). At $t = 0$, a primed-continuous ICV infusion of either insulin ($30 \mu\text{U}$ total dose), glibenclamide (K_{ATP} channel blocker, 3 nmol), diazoxide (1.5 nmol) or vehicle (aCSF or 5% dimethyl sulphoxide) was initiated and maintained for the remainder of the experiment (Fig. 1a). In separate experiments, insulin, glibenclamide or diazoxide was infused

intrahypothalamically with a similar time course, at the total dose of $2 \mu\text{U}$, 200 pmol, and 100 pmol, respectively. Pancreatic insulin-clamp studies were performed as previously described²⁵. Euglycaemic-hyperinsulinaemic clamps in conscious, unrestrained, catheterized mice were performed for 90 min as previously described²⁷.

Selective hepatic branch vagotomy

One week before the clamp procedure, a subgroup of rats underwent selective hepatic vagotomy (HV) or a sham operation (sham). Anatomical nerve transections were verified in each rat at sacrifice by microscopic observation of the absence of vagal nerve fibres between the two silk sutures that had been placed along the trunk at the time of transection²⁸.

Selective vagal deafferentation

Nerve transection was performed according to previously published methods²⁹.

Expression of gluconeogenic enzymes

Quantitative analysis of gene expression was done using quantitative PCR. Total RNA was isolated with Trizol (Invitrogen) and single-stranded complementary DNA was synthesized using a kit provided by Invitrogen. Real-time PCR reactions and the primers for *G6Pase* and *Pepck* were done as previously described²⁵. The copy number of each transcript was derived from a standard curve of cloned target templates. Expression of each transcript was normalized to the copy number for 18S ribosomal protein.

Detection of *Sur1* and *Sur2* mRNA in hypothalamic nuclei

PCR with primer sets specific for either rat *Sur1* (forward, 5′-CTTCCAGACAGCGAG GGAGAA-3′; reverse, 5′-TCTCCATGGGGCAGGATGTCT-3′) or *Sur2* (forward, 5′-ATGAGCCTTTCCTTCTGTGGT-3′; reverse, 5′-CGCCCGTTGCGAGTCTGAAACA-3′) was performed on cDNA synthesized from total RNA extracted from rat arcuate, lateral hypothalamus and paraventricular nucleus. Expression of *Sur1* and *Sur2* was compared to that of other tissues, including rat pancreatic islets, mouse pancreatic β -TC-3 cells and rat heart. The resulting PCR products were resolved on a 2% agarose gel and visualized after staining with ethidium bromide. The expected size for the PCR products is 230 base pairs (bp) for *Sur1* and 294 bp for *Sur2*.

Brain stereotactic ‘micro punches’

Brain ‘micro punches’ of individual hypothalamic nuclei were prepared as previously described⁶.

Hepatic glucose fluxes

The rates of hepatic glucose fluxes were determined as previously described^{25,27}.

Additional methods

All values are presented as means $\pm$ s.e.m. Comparisons among groups were made using analysis of variance (ANOVA) or unpaired Student’s *t* tests, as appropriate. The study protocol was reviewed and approved by the Institutional Animal Care and Use Committee of the Albert Einstein College of Medicine.

Received 22 November 2004; accepted 19 February 2005; doi:10.1038/nature03439.

- Friedman, J. M. A war on obesity, not the obese. *Science* **299**, 856–858 (2003).
- Flier, J. S. Obesity wars: molecular progress confronts an expanding epidemic. *Cell* **116**, 337–350 (2004).
- Magnusson, L., Rothman, D. L., Katz, L. D., Shulman, R. G. & Shulman, G. I. Increased rate of gluconeogenesis in type II diabetes mellitus. A ¹³C nuclear magnetic resonance study. *J. Clin. Invest.* **90**, 1323–1327 (1992).
- Schwartz, M. W., Woods, S. C., Porte, D. Jr, Seeley, R. J. & Baskin, D. G. Central nervous system control of food intake. *Nature* **404**, 661–671 (2000).
- Obici, S., Zhang, B. B., Karkanias, G. & Rossetti, L. Hypothalamic insulin signaling is required for inhibition of glucose production. *Nature Med.* **8**, 1376–1382 (2002).
- Obici, S., Feng, Z., Arduini, A., Conti, R. & Rossetti, L. Inhibition of hypothalamic carnitine palmitoyltransferase-1 decreases food intake and glucose production. *Nature Med.* **9**, 756–761 (2003).
- Obici, S., Feng, Z., Karkanias, G., Baskin, D. G. & Rossetti, L. Decreasing hypothalamic insulin receptors causes hyperphagia and insulin resistance in rats. *Nature Neurosci.* **5**, 566–572 (2002).
- Bruning, J. C. *et al.* Role of brain insulin receptor in control of body weight and reproduction. *Science* **289**, 2122–2125 (2000).
- Woods, S. C., Lotter, E. C., McKay, L. D. & Porte, D. Jr Chronic intracerebroventricular infusion of insulin reduces food intake and body weight of baboons. *Nature* **282**, 503–505 (1979).
- Niswender, K. D. *et al.* Insulin activation of phosphatidylinositol 3-kinase in the hypothalamic arcuate nucleus: a key mediator of insulin-induced anorexia. *Diabetes* **52**, 227–231 (2003).
- Aguilar-Bryan, L. & Bryan, J. Molecular biology of adenosine triphosphate-sensitive potassium channels. *Endocr. Rev.* **20**, 101–135 (1999).
- Seghers, V., Nakazaki, M., DeMayo, F., Aguilar-Bryan, L. & Bryan, J. *Sur1* knockout mice. A model for K_{ATP} channel-independent regulation of insulin secretion. *J. Biol. Chem.* **275**, 9270–9277 (2000).
- Karschin, C., Ecke, C., Ashcroft, F. M. & Karschin, A. Overlapping distribution of K_{ATP} channel-forming Kir6.2 subunit and the sulfonylurea receptor SUR1 in rodent brain. *FEBS Lett.* **401**, 59–64 (1997).
- Spaniswick, D., Smith, M. A., Groppi, V. E., Logan, S. D. & Ashford, M. L. Leptin inhibits hypothalamic neurons by activation of ATP-sensitive potassium channels. *Nature* **390**, 521–525 (1997).
- Spaniswick, D., Smith, M. A., Mirshamsi, S., Routh, V. H. & Ashford, M. L. Insulin activates ATP-sensitive K⁺ channels in hypothalamic neurons of lean, but not obese rats. *Nature Neurosci.* **3**, 757–758 (2000).
- Grill, H. J. *et al.* Evidence that the caudal brainstem is a target for the inhibitory effect of leptin on food intake. *Endocrinology* **143**, 239–246 (2002).

17. Aguilar-Bryan, L. *et al.* Cloning of the beta cell high-affinity sulfonylurea receptor: a regulator of insulin secretion. *Science* **268**, 423–426 (1995).
18. Seino, S. & Miki, T. Physiological and pathophysiological roles of ATP-sensitive K⁺ channels. *Prog. Biophys. Mol. Biol.* **81**, 133–176 (2003).
19. Seino, S., Iwanaga, T., Nagashima, K. & Miki, T. Diverse roles of K_{ATP} channels learned from Kir6.2 genetically engineered mice. *Diabetes* **49**, 311–318 (2000).
20. Chutkow, W. A. *et al.* Disruption of *Sur2*-containing K_{ATP} channels enhances insulin-stimulated glucose uptake in skeletal muscle. *Proc. Natl Acad. Sci. USA* **98**, 11760–11764 (2001).
21. Matsuhisa, M. *et al.* Important role of the hepatic vagus nerve in glucose uptake and production by the liver. *Metabolism* **49**, 11–16 (2000).
22. Loftus, T. M. *et al.* Reduced food intake and body weight in mice treated with fatty acid synthase inhibitors. *Science* **288**, 2379–2381 (2000).
23. Minokoshi, Y. *et al.* AMP-kinase regulates food intake by responding to hormonal and nutrient signals in the hypothalamus. *Nature* **428**, 569–574 (2004).
24. Obici, S. *et al.* Central administration of oleic acid inhibits glucose production and food intake. *Diabetes* **51**, 271–275 (2002).
25. Pocai, A., Obici, S., Schwartz, G. J. & Rossetti, L. A brain–liver circuit regulates glucose homeostasis. *Cell Metab.* **1**, 53–61 (2005).
26. Morton, G. J. *et al.* Arcuate nucleus-specific leptin receptor gene therapy attenuates the obesity phenotype of Koletsky (*fa^h/fa^h*) rats. *Endocrinology* **144**, 2016–2024 (2003).
27. Massillon, D. *et al.* Quantitation of hepatic glucose fluxes and pathways of hepatic glycogen synthesis in conscious mice. *Am. J. Physiol.* **269**, E1037–E1043 (1995).
28. la Fleur, S. E., Ji, H., Manalo, S. L., Friedman, M. I. & Dallman, M. F. The hepatic vagus mediates fat-induced inhibition of diabetic hyperphagia. *Diabetes* **52**, 2321–2330 (2003).
29. Norgren, R. & Smith, G. P. Central distribution of subdiaphragmatic vagal branches in the rat. *J. Comp. Neurol.* **273**, 207–223 (1988).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We wish to thank B. Liu, S. Gaweda and C. Baveghems for expert technical assistance. This work was supported by the NIH, ADA and the Skirball Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to L.R. (rossetti@acom.yu.edu).

Sox2 is required for sensory organ development in the mammalian inner ear

Amy E. Kiernan^{1,*,†}, Anna L. Pelling^{2,*}, Keith K. H. Leung², Anna S. P. Tang², Donald M. Bell³, Charles Tease^{4,†}, Robin Lovell-Badge³, Karen P. Steel^{1,†} & Kathryn S. E. Cheah²

¹MRC Institute of Hearing Research, University of Nottingham, Nottingham NG7 2RD, UK

²Department of Biochemistry, The University of Hong Kong, Faculty of Medicine Building, 21 Sassoon Road, Pokfulam, Hong Kong, China

³Division of Developmental Genetics, MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

⁴MRC Mammalian Genetics Unit, Harwell, Didcot OX11 0RD, UK

* These authors contributed equally to this work

† Present addresses: The Jackson Laboratory, 600 Main Street, Bar Harbor, Maine 04609, USA (A.E.K.); Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK (K.P.S.); Science Style, Courtland Road, Oxford OX4 4JB, UK (C.T.)

Sensory hair cells and their associated non-sensory supporting cells in the inner ear are fundamental for hearing and balance. They arise from a common progenitor¹, but little is known about the molecular events specifying this cell lineage. We recently identified two allelic mouse mutants, light coat and circling (*Lcc*) and yellow submarine (*Ysb*), that show hearing and balance impairment². *Lcc/Lcc* mice are completely deaf, whereas *Ysb/Ysb* mice are severely hearing impaired². We report here that inner ears of *Lcc/Lcc* mice fail to establish a prosensory domain and neither hair cells nor supporting cells differentiate, resulting in a severe inner ear malformation, whereas the sensory epithelium

of *Ysb/Ysb* mice shows abnormal development with disorganized and fewer hair cells. These phenotypes are due to the absence (in *Lcc* mutants) or reduced expression (in *Ysb* mutants) of the transcription factor SOX2, specifically within the developing inner ear. SOX2 continues to be expressed in the inner ears of mice lacking *Math1* (also known as *Atoh1* and *HATH1*), a gene essential for hair cell differentiation, whereas *Math1* expression is absent in *Lcc* mutants, suggesting that *Sox2* acts upstream of *Math1*.

In the mammalian inner ear, six functionally and morphologically distinct sensory regions are required for transducing sound and vestibular information. Five of these regions are important for balance whereas one, the organ of Corti, is required for hearing. Despite the diverse functions, each sensory region contains the same basic cell types: the sensory hair cells and associated supporting cells. Similar to *Drosophila* sense organ development, *Notch*-mediated lateral inhibition is required to generate the correct numbers of sensory hair cells³. In *Drosophila*, proneural clusters are first delineated by bHLH proteins such as members of the *achaete-scute* and *atonal* gene families⁴. Within the sensory areas of the mammalian inner ear, *Math1*, an orthologue of the *atonal* gene, is required for the differentiation or selection of hair cells but not for the initial specification of the sensory regions themselves^{5,6}. Moreover, nothing is known of the pathways upstream of *Math1* that specify sensory organ formation.

The two allelic mouse mutants *Ysb*, which arose owing to a fortuitous insertion of a transgene (pAA2)², and *Lcc*, caused by X-ray irradiation⁷, are both localized to chromosome 3 and involve complex chromosomal rearrangements and/or deletions^{2,7}. To understand the morphological defects that lead to the hearing and balance impairments, inner ears from both mutants were analysed by paint-filling at embryonic day (E)16.5. Both mutants showed abnormal inner ear morphology, but a more severe malformation was observed in inner ears of *Lcc/Lcc* mice (Fig. 1a–c). In *Lcc/Lcc* mice, all three ampullae, structures that house the sensory cristae, were missing and the semicircular canals showed only the most rudimentary development (Fig. 1c). In addition, the cochlea was undercoiled and the saccule and utricle were extremely small. Inner ears of *Ysb/Ysb* mice were less severely affected: both the anterior and lateral ampullae were missing and the respective semicircular canals were truncated (Fig. 1b). The cochlea was slightly undercoiled and sometimes appeared swollen. The inner ears of *Ysb/Lcc* compound heterozygotes displayed an intermediate phenotype (Fig. 1d).

Because the most severely affected regions of mutant inner ears were those that contained sensory regions, we examined whether sensory development was disrupted in *Lcc/Lcc* and *Ysb/Ysb* mutants. Sensory regions were examined at birth by scanning electron microscopy, and a severe disruption in sensory organ development was observed in both mutants (Fig. 1; see also Supplementary Fig. 1). No evidence of inner ear hair cell formation was observed in *Lcc/Lcc* mice at postnatal day (P)0 (Fig. 1i, l) and some sensory formation occurred in *Ysb/Ysb* mice, although hair cell patterning was clearly abnormal. The very basal regions of the cochlea contained patches of hair cells, with no clear delineation between inner and outer hair cell types (Fig. 1h, k) and hair cells in the apical region of the cochlea formed continuous but irregular rows (Supplementary Fig. 1b). The three vestibular organs that developed in *Ysb/Ysb* mice also demonstrated sensory defects, and the utricle was most severely affected, with almost no hair cell formation (Supplementary Fig. 1c–k). Despite the abnormal patterning, the cochlear hair cells that developed in *Ysb/Ysb* mice survived until at least P30 (data not shown) and showed some function².

Sox2, a member of the group B Sox (SRY-related HMG box) family, maps to the vicinity of the *Ysb* and *Lcc* chromosomal rearrangements and is expressed in the developing inner ear of both chicken and mouse^{8,9}. A complementation test was carried out

by crossing both the *Ysb* and *Lcc* alleles to a targeted null allele, *Sox2*^{*βgeo*} (ref. 10). Similar to *Ysb/Ysb* and *Lcc/Lcc* mutants², *Ysb/Sox2*^{*βgeo*} compound heterozygotes exhibited yellow coats as well as circling and deafness, indicating non-complementation. *Lcc/Sox2*^{*βgeo*} compound heterozygotes did not survive postnatally, but were recovered in normal mendelian frequencies at E16.5, allowing inner ear analyses. Paint-filling of compound heterozygote *Ysb/Sox2*^{*βgeo*} and *Lcc/Sox2*^{*βgeo*} inner ears (Fig. 1e, f) revealed a marked similarity to the defects observed in *Ysb/Lcc* and *Lcc/Lcc* inner ears, respectively. This identifies *Sox2* as the affected gene in both mutants.

Mouse embryos homozygous for *Sox2*^{*βgeo*} die shortly after implantation¹⁰, thus *Ysb* and *Lcc* do not represent null alleles. Southern hybridization analysis and polymerase chain reaction (PCR) amplification of genomic PCR showed that the *Sox2* coding region and nearby flanking DNA were intact in both *Ysb* and *Lcc*, whereas reverse transcription (RT)-PCR and northern hybridization analyses using whole-embryo RNA revealed normal *Sox2* transcripts (data not shown). Analyses of temporal and spatial expression patterns revealed that *Sox2* was expressed in the neural tube and otocyst of wild-type embryos at E9.5, with strong expression in the ventral and lateral regions of the otocyst, consistent with a role in sensory development^{11,12} (Fig. 2a, e). In *Lcc/Lcc*

embryos, however, although *Sox2* expression appeared normal throughout most of the neural tube, it was not detected in the otocyst or adjacent region of the hindbrain (Fig. 2c, g). *Sox2* expression was also significantly reduced in *Ysb/Ysb* otocysts (Fig. 2b, f).

Regulation of *Sox2* expression involves multiple separate tissue-specific enhancers^{13,14}. The genetic complementation results and expression data indicate that *Lcc* and *Ysb* are both regulatory alleles of *Sox2* in which otic-specific transcription is compromised. The *Lcc* mutation leads to complete loss of *Sox2* expression in the otocyst, possibly through loss or translocation of *Sox2* tissue-specific regulatory elements. The partial loss of *Sox2* expression in *Ysb* mice could also be explained by regulatory element loss or translocation but is more likely due to the juxtaposition of *Sox2* regulatory elements and the *Ysb* transgenic reporter construct (pAA2)², resulting in promoter/enhancer interference or competition. The pAA2 transgene consists of regulatory sequences of the *COL2A1* gene linked to a *lacZ* reporter², the expression of which is usually restricted to a subset of *Col2a1* expression domains at E9.5 in transgenic mice, such as the notochord, somites and frontonasal mesenchyme¹⁵. However, in +/*Ysb* embryos, additional sites of *lacZ* expression were observed in regions that overlapped with sites of *Sox2* expression, including the ventral otic vesicle and regions of the

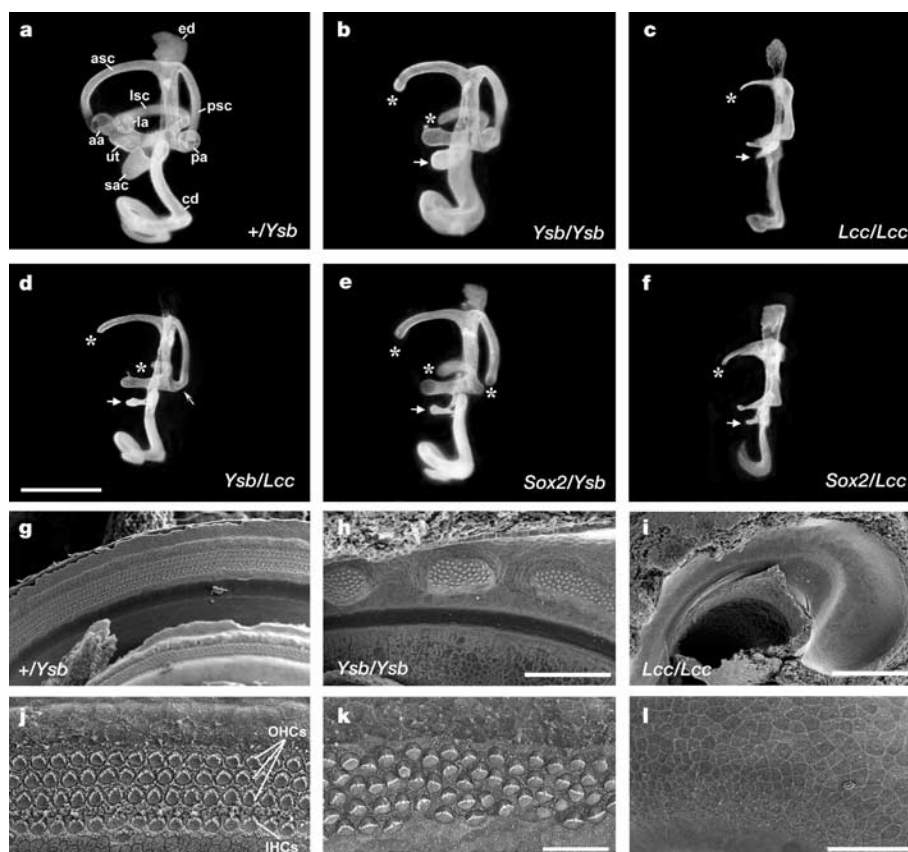


Figure 1 Inner ear malformations and sensory defects in *Ysb* and *Lcc* mice. **a–f**, Paint-filled inner ears at E16.5 from *Ysb/Ysb* ($n=8$ mutants; 4 controls), *Lcc/Lcc* ($n=10$ mutants; 5 controls) and compound heterozygote mutant mice ($n=6$ *Ysb/Lcc*; 6 *Ysb/Sox2*^{*βgeo*}; 4 *Lcc/Sox2*^{*βgeo*}; 10 controls). Asterisks indicate truncated semicircular canals and missing ampullae in mutants. Large arrows in **b–f** point to the saccule, which is present but shows varying degrees of malformation in the different allelic combinations. The small arrow in **d** indicates the smaller posterior ampulla, which is sometimes absent in the inner ears of *Ysb/Lcc* mice. The endolymphatic duct (ed) is present in **b** and **d** but not filled well in these particular examples. **g–l**, Scanning electron micrographs depicting organ of Corti basal portions at P0 from *Ysb/Ysb* ($n=8$

mutants; 5 controls) and *Lcc/Lcc* ($n=4$ mutants; 2 controls) mice. Note the abnormal patches of hair cells found in *Ysb/Ysb* mice (**h, k**) whereas no sensory hair cells were detected in *Lcc/Lcc* mice (**i, l**). Inner and outer hair cells (IHCs and OHCs, respectively) could be distinguished in control cochleae (**j**) whereas no such distinction could be made in the cochleae of *Ysb/Ysb* mice (**k**). aa, anterior ampulla; asc, anterior semicircular canal; cd, cochlear duct; ed, endolymphatic duct; la, lateral ampulla; lsc, lateral semicircular canal; pa, posterior ampulla; psc, posterior semicircular canal; sac, saccule; ut, utricle. Scale bar: 1 mm (**a–f**), 100 μ m (**g, h**), 200 μ m (**i, j**) and 50 μ m (**k, l**).

neural tube (Fig. 2d, h, and data not shown). *Lcc* and *Ysb* therefore represent unique alleles of *Sox2* in which complex chromosomal rearrangements have resulted in the loss of, or interference with, specific regulatory elements that direct expression of *Sox2* within the inner ear.

Sox2 marks neural progenitor cells at different stages of central nervous system development¹⁶, and evidence from *Drosophila* and *Xenopus* suggests that group B1 *Sox* genes are required for the specification or consolidation of early neural fate^{17–19}. Moreover, early in development *Sox2* is essential for specifying the ICM/epiblast cell lineage¹⁰. To investigate whether SOX2 has a role in specifying or maintaining the sensory progenitors in the inner ear, we examined expression of the SOX2 protein at E14.5 and compared it with that of P27KIP1 (also known as CDKN1B or KIP1), an established marker of the prosensory domain in the cochlea²⁰. Immunostaining of wild-type embryos revealed that the SOX2 protein was expressed uniformly throughout the epithelium of all six sensory regions (Fig. 2i) and demonstrated an overlapping expression pattern with P27KIP1 (Fig. 2m), consistent with an early role in inner ear sensory development. Similar to E9.5, SOX2 protein was not detected in *Lcc/Lcc* inner ears at E14.5, and was reduced in the inner ears of *Ysb/Ysb* mice (Fig. 2j, k). P27KIP1 was also absent in *Lcc/Lcc* cochleae (Fig. 2o), indicating a failure in the establishment of the prosensory domain. The sensory epithelium does develop in the cochlea of *Ysb/Ysb* mice, so as expected, P27KIP1 expression was detected in *Ysb/Ysb* inner ears at E14.5 (Fig. 2n). There was no evidence for premature hair cell differentiation in either *Ysb* or *Lcc* mutants at E14.5 (data not shown), which argues against a role for SOX2 in maintaining or prolonging stem cell status in the ear at this stage of development.

By E15.5/E16.5, cells in the wild-type organ of Corti are differentiating into hair cells and supporting cells, and SOX2 was detected in both cell types in control inner ears (Fig. 3a, j). SOX2 expression remained undetectable in *Lcc/Lcc* inner ears at these stages (Fig. 3c, l; Supplementary Fig. 2i, l) and, consistent with the absence of the prosensory domain, early markers of hair and supporting cell differentiation, including MYO7A, P27KIP1 and *Lfng*, were also absent (Fig. 3f, i, l, o, r, u; see also Supplementary Fig. 2c, f). In the *Ysb/Ysb* organ of Corti, MYO7A, P27KIP1 and *Lfng* were expressed (Fig. 3e, h, n, q, t), as were later markers for hair cell differentiation such as *Pou4f3* and GF11 (Supplementary Fig. 3), but all showed abnormal expression patterns consistent with the irregular pattern of hair cells observed at birth. At E15.5/16.5, expression of the SOX2 protein was markedly downregulated in *Ysb/Ysb* hair cells, although expression in supporting cells appeared relatively normal (Fig. 3b, k). This differential effect of the *Ysb* mutation on SOX2 expression in the hair and supporting cells may contribute to the disorganized pattern of hair cell differentiation in the organ of Corti of *Ysb/Ysb* mice. In the *Ysb/Ysb* vestibule, SOX2 expression was also reduced and the sensory regions that were present (including the utricular and saccular maculae and the posterior cristae) contained fewer hair cells than wild type (Supplementary Figs 1 and 2b, e, h, k). This was particularly evident in the utricle, where SOX2 expression was barely detectable and almost no hair cells were found (Supplementary Fig. 2e, k). Given the continued expression observed at E16.5 in hair cells and supporting cells there may also be a later role for SOX2 in differentiated sensory and/or non-sensory supporting cells.

Math1 is one of the earliest identified markers of hair cell differentiation and is essential for their genesis^{5,21}. It was initially proposed that *Math1* may be the factor required for prosensory

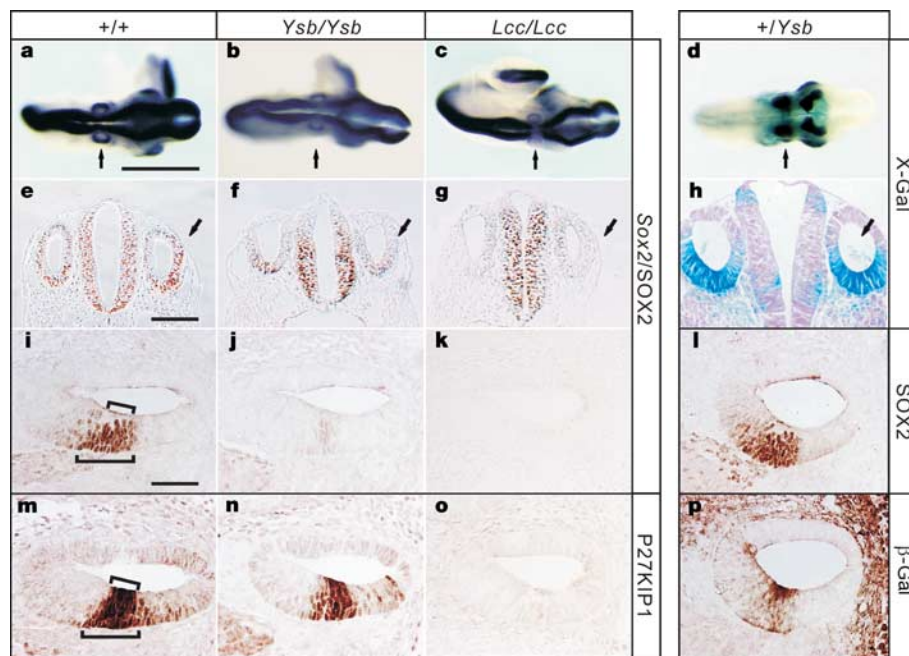


Figure 2 Altered expression of the *Sox2* gene and protein in *Ysb* and *Lcc* mice and loss of inner ear prosensory domains. **a–c**, Whole-mount *in situ* hybridization at E9.5. Arrows indicate lack of *Sox2* in the *Lcc/Lcc* otocyst and adjacent hindbrain and *Sox2* downregulation in the otocyst of *Ysb/Ysb* embryos. **d**, Ectopic β -galactosidase expression in the neural tube and otic vesicle (arrow) in *+Ysb* mice at E9.5. **e–g**, Immunohistochemistry of otocyst sections at E9. Arrows indicate that SOX2 is not detected in the *Lcc/Lcc* otocyst and reduced in that of *Ysb/Ysb* mice. **h**, Arrow indicates β -galactosidase expression in the *+Ysb* otocysts at E9.5. **i–o**, Immunohistochemistry of E14.5 cochlear sections using either an antibody to SOX2 (**i–l**) or P27KIP1 (**m–o**). Brackets in **i**

indicate the SOX2 expression domain that overlaps with that of P27KIP1 (**m**), which marks the nascent organ of Corti. SOX2 expression is reduced in the organ of Corti in *Ysb/Ysb* mice (**j**) and absent in *Lcc/Lcc* mice (**k**). P27KIP1 is present in the cochleae of *Ysb/Ysb* mice (**n**) but absent in *Lcc/Lcc* mice (**o**). **p**, Immunohistochemistry using an antibody to detect β -galactosidase in *+Ysb* mice demonstrates a similar domain of transgene expression to that of SOX2 in the organ of Corti (**l**). Mesenchymal expression in **p** matches normal expression of the pAA2 transgene. Scale bar: 1 mm (**a–d**), 100 μ m (**e–h**) and 50 μ m (**i–p**).

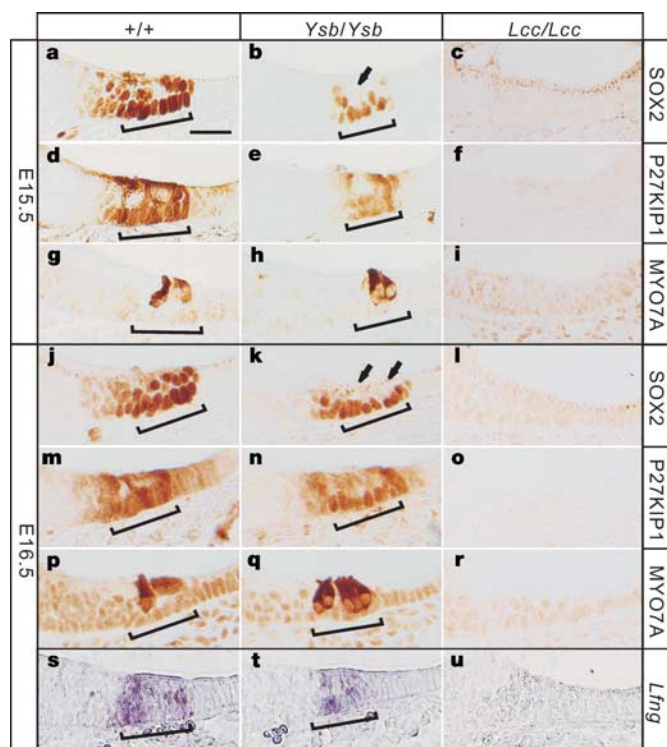


Figure 3 Altered or loss of expression of SOX2 protein correlates with abnormal/loss of hair and supporting cell differentiation in *Ysb/Ysb* and *Lcc/Lcc* mice, respectively. **a–r**, Immunohistochemistry of E15.5 (**a–i**) and E16.5 (**j–r**) cochlear sections using antibodies to SOX2 (**a–c**, **j–l**), P27KIP1 (**d–f**, **m–o**) and MYO7A (**g–i**, **p–r**). **s–u**, *In situ* hybridization using a *Lfng*-specific riboprobe. Cochleae from *Lcc/Lcc* mice show loss of expression of SOX2 and hair (MYO7A) and supporting cell (P27KIP1 and *Lfng*) markers. Cochleae from *Ysb/Ysb* mice show abnormal hair/supporting cell differentiation and strong SOX2 downregulation in hair cells. Brackets indicate sensory epithelium and arrows indicate position of hair cells. Scale bar, 50 μ m.

establishment in the inner ear. However, it has been demonstrated that *Math1* is not required for the genesis of supporting cells in the inner ear and therefore is unlikely to be required for prosensory specification^{6,22}. Unlike the inner ears of *Lcc/Lcc* mice, P27KIP1 is expressed normally in *Math1* null mutants⁶, further confirming that *Math1* is not a prosensory gene. *Math1* transcripts were not detected in *Lcc/Lcc* inner ears at E15.5 (Fig. 4b), indicating that hair cell differentiation is never initiated in these mutants. As expected, *Math1* expression was detected in *Ysb/Ysb* mice (Fig. 4d). The SOX2 protein was also expressed in *Math1* null mutant inner ears at E14.5 (Fig. 4f), which indicates that *Sox2* is likely to be acting upstream of *Math1* during sensory development in the inner ear, consistent with a role in prosensory establishment.

In addition to sensory defects, both *Lcc* and to a lesser degree *Ysb* mice display defects in non-sensory regions of the ear, such as the semicircular canals and non-sensory regions of the cochlea and vestibule. Although it is possible that these defects are due directly to loss of *Sox2*, it is more likely that these defects are caused secondarily to loss of sensory development. Recently it has been shown that fibroblast growth factors expressed by the sensory cristae direct outgrowth of the canals by regulating BMP2 (ref. 23), providing molecular evidence that sensory areas can direct non-sensory development in the inner ear. Furthermore, the most abnormal non-sensory regions of the ear in both *Lcc* and *Ysb* mutant mice are those associated with the development of sensory structures, suggesting a secondary rather than primary effect. The severe malformations found in *Lcc/Lcc* mutant inner ears suggest it is

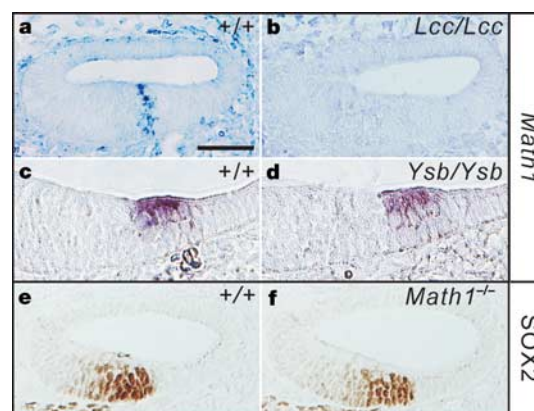


Figure 4 *Sox2* functions upstream of *Math1* during sensory organ formation. **a–d**, *In situ* hybridization of E15.5 (**a**, **b**) and E16.5 (**c**, **d**) cochlear sections shows no detectable *Math1* transcription in *Lcc/Lcc* mutants (**b**) and *Math1* expression in *Ysb/Ysb* hair cells (**d**). **e**, **f**, Immunohistochemistry of E14.5 cochlear sections demonstrates normal SOX2 expression in *Math1*^{−/−} mutants (**f**). Scale bar, 50 μ m.

likely that all the sensory regions have some influence on non-sensory development. Consistent with this hypothesis, structures that do not normally contain sensory regions, such as the common crus and endolymphatic duct and sac, developed relatively normally in *Lcc/Lcc* mice. Both *Lcc/Lcc* and *Ysb/Ysb* mice also displayed positive endocochlear potentials required for normal hair cell function (Supplementary Table 1), indicating that the non-sensory regions required to generate this potential developed normally in most cases.

A number of genes have previously been implicated in sensory region development of the inner ear^{24–28}; however, none has been shown to be essential for the development of all six sensory regions. Thus, it was unclear whether these regions develop from a common origin using the same genetic pathway, or are established independently. Our results indicate a common molecule underlying sensory establishment in the ear. Furthermore, we propose a genetic hierarchy in which *Sox2* functions upstream of *Math1* to establish the sensory progenitor cells.

Unfortunately, hair cells in the adult mammalian cochlea cannot regenerate, unlike those in some other vertebrates. Both embryonic stem cells and stem cells derived from the adult mouse utricle have recently been shown to be capable of differentiating into hair-cell-like cells^{29,30}. *Sox2* has been shown to be a distinguishing marker of neural stem cells¹⁶, and our data demonstrate the importance of its role in establishing sensory progenitors in the ear. It will be interesting to determine whether *Sox2* is important for the establishment and/or maintenance of this stem cell type and whether *Sox2* could have a function in sensory patch regeneration. □

Methods

Mice and genotyping

Lcc and *Ysb* mutant mouse strains were previously described². All experiments were authorized by licences from the Hong Kong Government Department of Health, University of Hong Kong Committee on the Use of Live Animals in Teaching and Research, and the UK Home Office. *Math1*^{βGal} mice were supplied by H. Zoghbi⁵. *Ysb* mice were genotyped according to ref. 2. *Lcc* mice were genotyped by PCR amplification of a closely linked natural polymorphism (Supplementary Fig. 4).

Paint-filling and scanning electron microscopy

The lumen of the inner ears of E16.5 embryos were filled with white paint as described previously²⁶. Scanning electron microscopy was performed using the OTOTO method²⁶.

Immunohistochemistry, *in situ* hybridization and X-gal staining

Wild-type and mutant embryos were fixed and processed using standard procedures^{10,15}. Immunohistochemistry was carried out using the Envision+ System (DAKO) and

HRP-DAB colorimetric detection. Primary antibodies used were specific for: SOX2 (ref. 10); β -galactosidase (ICN Pharmaceuticals); MYO7A (gift from A. El-Amraoui and C. Petit); P27KIP1 (NeoMarkers); and GF11 (gift from H. Bellen). X-Gal staining was as previously described¹⁵. Whole-mount and section *in situ* hybridizations were carried out according to standard procedures with digoxigenin-labelled riboprobes specific for Sox2 (ref. 8), *Math1* (gift from M. Kelley and R. Kageyama), *Lfng*²⁸ and *Pou4f3* (gift from K. Avraham). Anti-dig-AP antibody and BM purple (Roche) or NBT/BCIP colorimetric signal detection were used for whole-mount and section *in situ* hybridizations, respectively. Tyramide signal amplification (TSA Biotin System, PerkinElmer Life Sciences) was used for amplified section *in situ* hybridization.

Electrophysiology

Mice were anaesthetized with urethane, the cochlear wall was exposed, and a micropipette electrode inserted into scala media to measure the endocochlear potential²⁶.

Received 16 December 2004; accepted 7 February 2005; doi:10.1038/nature03487.

1. Fekete, D. M., Muthukumar, S. & Karagozeos, D. Hair cells and supporting cells share a common progenitor in the avian inner ear. *J. Neurosci.* **18**, 7811–7821 (1998).
2. Dong, S. *et al.* Circling, deafness, and yellow coat displayed by yellow submarine (*Ysb*) and light coat and circling (*Lcc*) mice with mutations on chromosome 3. *Genomics* **79**, 777–784 (2002).
3. Eddison, M., Le, R. I. & Lewis, J. Notch signaling in the development of the inner ear: lessons from *Drosophila*. *Proc. Natl Acad. Sci. USA* **97**, 11692–11699 (2000).
4. Jan, Y. N. & Jan, L. Y. Neuronal cell fate specification in *Drosophila*. *Curr. Opin. Neurobiol.* **4**, 8–13 (1994).
5. Bermingham, N. A. *et al.* *Math1*: an essential gene for the generation of inner ear hair cells. *Science* **284**, 1837–1841 (1999).
6. Chen, P., Johnson, J. E., Zoghbi, H. Y. & Segil, N. The role of *Math1* in inner ear development: Uncoupling the establishment of the sensory primordium from hair cell fate determination. *Development* **129**, 2495–2505 (2002).
7. Lyon, M. F., Phillips, R. J. & Fisher, G. Dose-response curves for radiation-induced gene mutations in mouse oocytes and their interpretation. *Mutat. Res.* **63**, 161–173 (1979).
8. Wood, H. B. & Episkopou, V. Comparative expression of the mouse *Sox1*, *Sox2* and *Sox3* genes from pre-gastrulation to early somite stages. *Mech. Dev.* **86**, 197–201 (1999).
9. Uchikawa, M., Kamachi, Y. & Kondoh, H. Two distinct subgroups of Group B Sox genes for transcriptional activators and repressors: their expression during embryonic organogenesis of the chicken. *Mech. Dev.* **84**, 103–120 (1999).
10. Avilion, A. A. *et al.* Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17**, 126–140 (2003).
11. Knowlton, V. Y. Correlation of the development of membranous and bony labyrinths, acoustics ganglia, nerves, and brain centers of the chick embryos. *J. Morphol.* **121**, 179–208 (1967).
12. Cole, L. K. *et al.* Sensory organ generation in the chicken inner ear: contributions of *bone morphogenetic protein 4*, *serrate1*, and *lunatic fringe*. *J. Comp. Neurol.* **424**, 509–520 (2000).
13. Zappone, M. V. *et al.* Sox2 regulatory sequences direct expression of a (beta)-geo transgene to telencephalic neural stem cells and precursors of the mouse embryo, revealing regionalization of gene expression in CNS stem cells. *Development* **127**, 2367–2382 (2000).
14. Uchikawa, M., Ishida, Y., Takemoto, T., Kamachi, Y. & Kondoh, H. Functional analysis of chicken *Sox2* enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Dev. Cell* **4**, 509–519 (2003).
15. Cheah, K. S. *et al.* Human COL2A1-directed SV40 T antigen expression in transgenic and chimeric mice results in abnormal skeletal development. *J. Cell Biol.* **128**, 223–237 (1995).
16. Cai, J. *et al.* Properties of a fetal multipotent neural stem cell (NEP cell). *Dev. Biol.* **251**, 221–240 (2002).
17. Buescher, M., Hing, F. S. & Chia, W. Formation of neuroblasts in the embryonic central nervous system of *Drosophila melanogaster* is controlled by *SoxNeuro*. *Development* **129**, 4193–4203 (2002).
18. Overton, P. M., Meadows, L. A., Urban, J. & Russell, S. Evidence for differential and redundant function of the Sox genes *Dichaete* and *SoxN* during CNS development in *Drosophila*. *Development* **129**, 4219–4228 (2002).
19. Kishi, M. *et al.* Requirement of *Sox2*-mediated signaling for differentiation of early *Xenopus* neuroectoderm. *Development* **127**, 791–800 (2000).
20. Chen, P. & Segil, N. p27^(Kip1) links cell proliferation to morphogenesis in the developing organ of Corti. *Development* **126**, 1581–1590 (1999).
21. Zheng, J. L. & Gao, W. Q. Overexpression of *Math1* induces robust production of extra hair cells in postnatal rat inner ears. *Nature Neurosci.* **3**, 580–586 (2000).
22. Woods, C., Montcouquiol, M. & Kelley, M. W. *Math1* regulates development of the sensory epithelium in the mammalian cochlea. *Nature Neurosci.* **7**, 1310–1318 (2004).
23. Chang, W., Brigande, J. V., Fekete, D. M. & Wu, D. K. The development of semicircular canals in the inner ear: role of FGFs in sensory cristae. *Development* **131**, 4201–4211 (2004).
24. Pirvola, U. *et al.* FGFR1 is required for the development of the auditory sensory epithelium. *Neuron* **35**, 671–680 (2002).
25. Pauley, S. *et al.* Expression and function of FGF10 in mammalian inner ear development. *Dev. Dyn.* **227**, 203–215 (2003).
26. Kiernan, A. E. *et al.* The Notch ligand *Jagged1* is required for inner ear sensory development. *Proc. Natl Acad. Sci. USA* **98**, 3873–3878 (2001).
27. Tsai, H. *et al.* The mouse *slalom* mutant demonstrates a role for *Jagged1* in neuroepithelial patterning in the organ of Corti. *Mol. Genet.* **10**, 507–512 (2001).
28. Morsli, H., Choo, D., Ryan, A., Johnson, R. & Wu, D. K. Development of the mouse inner ear and origin of its sensory organs. *J. Neurosci.* **18**, 3327–3335 (1998).
29. Li, H., Liu, H. & Heller, S. Pluripotent stem cells from the adult mouse inner ear. *Nature Med.* **9**, 1293–1299 (2003).
30. Li, H., Roblin, G., Liu, H. & Heller, S. Generation of hair cells by stepwise differentiation of embryonic stem cells. *Proc. Natl Acad. Sci. USA* **100**, 13495–13500 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements This work was supported by grants from the Research Grants Council (Hong Kong) to K.S.E.C., by the UK Medical Research Council, Defeating Deafness, the European Commission and the Louis Jeantet Foundation. We thank C. L. So and M. Chan for mouse care and research assistance, and H. Zoghbi for the gift of *Math1* null mice. We are also grateful to P. Tam, D. Loebel, V. Episkopou, J. Brigande and the RLB laboratory for critical comments and discussion.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to K.S.E.C. (hrmbdkc@hkusua.hku.hk).

Spatiotemporal regulation of MyD88–IRF-7 signalling for robust type-I interferon induction

Kenya Honda^{1*}, Yusuke Ohba^{1,2*}, Hideyuki Yanai^{1*}, Hideo Negishi¹, Tatsuaki Mizutani¹, Akinori Takaoka¹, Choji Taya³ & Tadatsugu Taniguchi¹

¹Department of Immunology, Graduate School of Medicine and Faculty of Medicine, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan

²Information and Cell Function, PRESTO, JST, Kawaguchi, Saitama 332-0012, Japan

³Department of Laboratory Animal Science, Tokyo Metropolitan Institute of Medical Science, Honkomagome 3-18-22, Bunkyo-ku, Tokyo 113-8613, Japan

* These authors contributed equally to this work

Robust type-I interferon (IFN- α/β) induction in plasmacytoid dendritic cells, through the activation of Toll-like receptor 9 (TLR9), constitutes a critical aspect of immunity^{1–6}. It is absolutely dependent on the transcription factor IRF-7, which interacts with and is activated by the adaptor MyD88 (ref. 7). How plasmacytoid dendritic cells, but not other cell types (such as conventional dendritic cells), are able to activate the MyD88–IRF-7-dependent IFN induction pathway remains unknown. Here we show that the spatiotemporal regulation of MyD88–IRF-7 signalling is critical for a high-level IFN induction in response to TLR9 activation. The IFN-inducing TLR9 ligand, A/D-type CpG oligodeoxynucleotide (CpG-A)^{3,4,8–11}, is retained for long periods in the endosomal vesicles of plasmacytoid dendritic cells, together with the MyD88–IRF-7 complex. However, in conventional dendritic cells, CpG-A is rapidly transferred to lysosomal vesicles. We further show that conventional dendritic cells can also mount a robust IFN induction if CpG-A is manipulated for endosomal retention using a cationic lipid. This strategy also allows us to demonstrate endosomal activation of the IFN pathway by the otherwise inactive TLR9 ligand B/K-type oligodeoxynucleotide (CpG-B)^{3,4,8–12}. Thus, our study offers insights into the regulation of TLR9 signalling in space, potentially suggesting a new avenue for therapeutic intervention.

Plasmacytoid dendritic cells (pDCs) represent a specialized cell population that is capable of producing large amounts of IFN- α/β ; therefore, they are also referred to as IFN-producing cells¹. Indeed, the activation of the IFN- α/β induction pathway in pDCs by the TLR9 subfamily (TLR7, TLR8 and TLR9) constitutes a critical aspect of immunity both in health and disease^{1–6}. This robust IFN induction in pDCs is absolutely dependent on the MyD88–IRF-7 signalling pathway⁷. These findings raise interesting but unanswered

questions of why and how pDCs, but not other cell types, activate this signalling pathway for IFN induction. The current conventional wisdom is that pDCs express a higher level of the transcription factor IRF-7 than other cell types^{11,13,14}, but it is unclear whether this accounts for the cell-type-specific, robust IFN induction.

Two classes of synthetic oligodeoxynucleotides (ODNs) containing an unmethylated CpG motif, CpG-A (also referred to as D-type ODN) and CpG-B (also referred to as K-type ODN), both activate TLR9 but elicit different responses. CpG-A induces IFN- α production much more efficiently than CpG-B in pDCs^{3,4,8–11} (Supplementary Fig. S1a). Using confocal microscopy, we first studied the intracellular trafficking of Cy5-labelled CpG-A (CpG-A-Cy5) and Cy5-labelled CpG-B (CpG-B-Cy5) in pDCs and conventional dendritic cells (cDCs) purified from *in vitro* cultured bone marrow cells or splenocytes. The ODNs D19 (ref. 8) and ODN1668 (ref. 12) were chosen as representatives of the CpG-A and CpG-B classes, respectively. In pDCs, CpG-A-Cy5 merged with fluorescein isothiocyanate (FITC)-labelled dextran, an endosomal marker, 90 min after its delivery; however, it did not merge with the lysosomal marker LysoTracker even 5 h after delivery (Fig. 1a and Supplementary Fig. S1b). In contrast, CpG-B-Cy5 merged with LysoTracker, but not with dextran-FITC, 90 min after its delivery (Fig. 1b). When CpG-B-Cy5 and FITC-labelled CpG-A were delivered together in pDCs, CpG-B-Cy5 (and not CpG-A-FITC) merged predominantly with LysoTracker (Fig. 1c). In cDCs, however, both CpG-A-Cy5 and CpG-B-Cy5 merged with LysoTracker 90 min after ODN delivery (Fig. 1d), an observation consistent with a previous report¹⁵. In pDCs, CpG-A-Cy5 also merged with yellow fluorescent protein

(YFP)-tagged IRF-7 (IRF-7-YFP) expressed by lentivirus-mediated gene transfer, but it failed to do so with IRF-7-YFP expressed in cDCs (Fig. 1e and Supplementary Fig. S1c). These observations suggest a model for spatiotemporal regulation of TLR9 signalling: CpG-A (but not CpG-B) bound to TLR9 remains in endosomal vesicles of pDCs to sustain the activation of the MyD88-IRF-7 pathway for IFN gene induction. Consistent with this model is the finding that the induction of both IRF-7 and IFN- α messenger RNAs continues even 12 h after CpG-A stimulation in pDCs and is dependent on *de novo* protein synthesis (Supplementary Fig. S1d).

To gain insight into the spatial regulation of CpG ODNs, MyD88 and IRF-7, we used the macrophage cell line RAW264.7, which readily permits the coexpression of various complementary DNAs. First, IRF-7-YFP or IRF-3-YFP was transiently coexpressed with cyan fluorescent protein (CFP)-tagged MyD88 (MyD88-CFP). IRF-7-YFP was expressed in the cytoplasm and merged with MyD88-CFP to form granular structures (Fig. 2a), whereas IRF-3-YFP failed to do so (Supplementary Fig. S2a). In addition, fluorescence resonance energy transfer (FRET) was preferentially observed between IRF-7-YFP and MyD88-CFP (Supplementary Fig. S2b), indicating a physical interaction between IRF-7 and MyD88. We found that IRF-7 interacted with MyD88 in the granular structures of several established cell lines, irrespective of the difference in its expression level or its localization pattern (refs 16, 17 and data not shown). Notably, the structures merged (albeit not exclusively) with fluorescently labelled dextran (Fig. 2a) but not with LysoTracker (Supplementary Fig. S2c). They could also be stained with

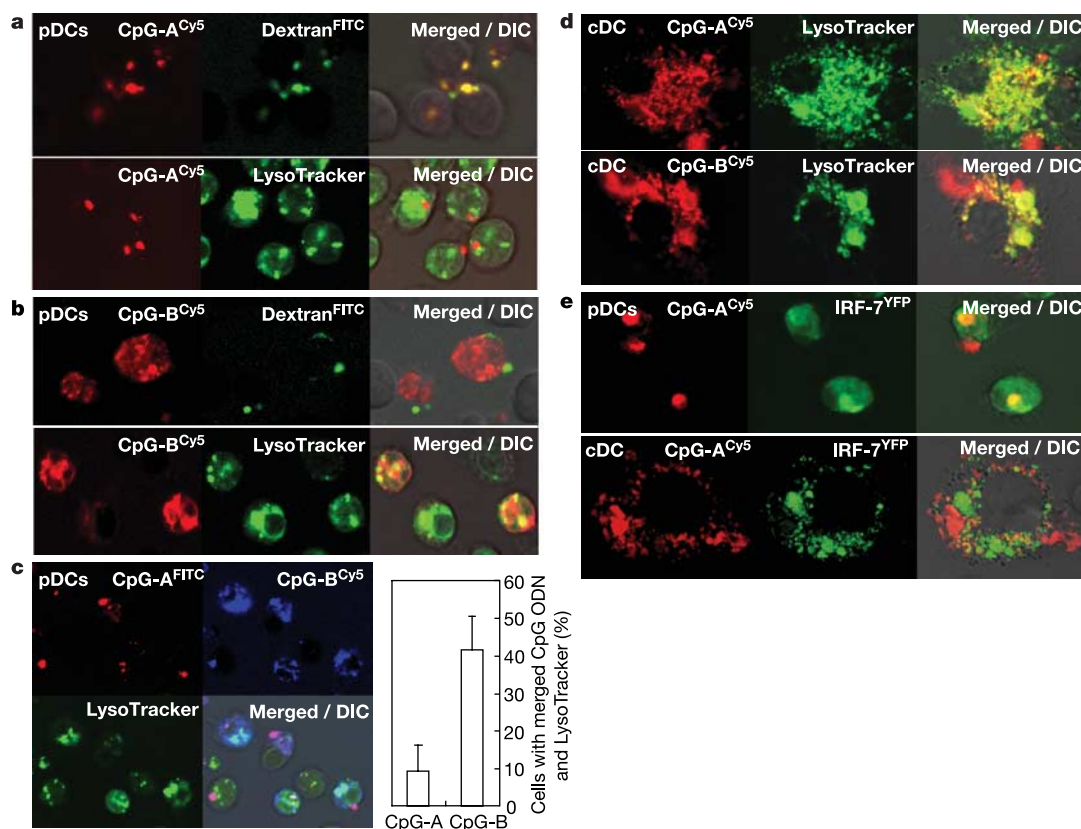


Figure 1 Trafficking of CpG ODNs in DCs. **a, b**, Flt3L-cultured bone-marrow (BM)-derived pDCs (120G8⁺ cells) were incubated with 3 μ M CpG-A-Cy5 (**a**) or CpG-B-Cy5 (**b**) for 90 min. To visualize endosomes or lysosomes, the cells were incubated with dextran-FITC or LysoTracker for the last 10 min in culture. Confocal and differential interference contrast (DIC) images of representative cells are shown in Figs 1–4. **c**, Splenic pDCs were

simultaneously incubated with 3 μ M CpG-A-FITC and CpG-B-Cy5, followed by incubation with LysoTracker. In the right panel, the percentage of cells in which CpG ODNs merged with LysoTracker is shown (error bars are s.d.). **d**, BM-derived cDCs (120G8[–] cells) were incubated with CpG ODNs and LysoTracker. **e**, BM-derived pDCs and cDCs expressing IRF-7-YFP were incubated with 3 μ M CpG-A-Cy5 for 90 min.

an anti-transferrin receptor antibody (Supplementary Fig. S2d), further indicating that the MyD88–IRF-7 complex resides in endosomal vesicles. On the other hand, when CpG-A–Cy5 was delivered into these cells, it merged with LysoTracker as early as 15 min after delivery (Fig. 2b and Supplementary Fig. S2e). Similar observations were made with CpG-B–Cy5 (Supplementary Fig. S2e). Thus, the pattern of trafficking of these ODNs in RAW cells is similar to that observed in cDCs. CpG-A–Cy5 did not merge with the MyD88–IRF-7 complex (Fig. 2c), and, as expected, the robust IFN- α/β mRNA induction by CpG-A did not occur in RAW cells (see below). These results are consistent with the above model, in that persistent TLR9 signalling and activation of the MyD88–IRF-7 pathway in endosomal vesicles are critical to activation of the robust IFN gene induction pathway.

As another approach to verifying the model, we asked whether a high IFN induction level could be achieved even in cDCs by manipulating the endosomal trafficking of CpG-A in order to resemble trafficking in pDCs. It has been reported that a complex consisting of DNA and cationic lipids can enter the cell through the endocytic pathway¹⁸. When cDCs were incubated with fluorescently labelled CpG-A complexed with the cationic lipid 1,2-dioleoyloxy-3-trimethylammonium-propane (DOTAP) (CpG-A–DOTAP), the complex merged with neither LysoTracker nor Lamp-1, another

lysosomal marker (Supplementary Videos 1–3 and Supplementary Fig. S3a). Instead, merging was seen with dextran, even 90 min after its delivery (Fig. 3a). Furthermore, CpG-A–Cy5–DOTAP merged with IRF-7–YFP expressed by the lentivirus in cDCs, and a substantial fraction of IRF-7–YFP was found in the nuclei of cDCs (Fig. 3b), indicating that the TLR9–MyD88–IRF-7 pathway has been activated in the endosomal vesicles of cDCs using this ODN delivery system.

Notably, a high level of IFN- α production (similar to that observed in pDCs) was found in cDCs stimulated with CpG-A–DOTAP (Fig. 3c and Supplementary Fig. S3b), despite their lower basal expression levels of IRF-7 compared with pDCs^{7,11,13,14}. This IFN- α expression in cDCs was accompanied by the efficient induction of IRF-7 mRNA (Supplementary Fig. S3c). DOTAP did not have an additive effect on IFN- α induction by CpG-A in pDCs (Fig. 3c and Supplementary Fig. S3b), suggesting that efficient endosomal signalling by CpG-A alone can activate the IFN induction pathway in pDCs. In both cell types, this induction was not observed when stimulated with a control ODN that has the CG dimer reversed within the CpG-A sequence (GpC ODN) and fails to activate TLR9 (refs 8, 10 and Supplementary Fig. S3b).

The IFN induction using both CpG-A stimulation regimens is TLR9-dependent, but partial IFN induction by CpG-A–DOTAP was observed in both pDCs and cDCs from TLR9 knockout (*Tlr9*^{−/−}) mice, indicating that an additional receptor must be activated by CpG-A–DOTAP (Fig. 3c and Supplementary Fig. S5). TLR7, the TLR9 subfamily member that is activated by single-stranded (ss)RNA^{19,20}, is also expressed in endosomes of cDCs and pDCs^{20,21}, and it is possible that TLR7 crossreacts with the CpG-A–DOTAP complex. Whatever the nature of the receptor, IFN- α induction was completely abolished in *MyD88*- and *Irf7*-null DCs (Fig. 3d), and was strongly inhibited by chloroquine or bafilomycin, both of which inhibit endosomal signalling²² (Fig. 3e). All these results support our endosomal trafficking model of CpG-A for activation of the MyD88–IRF-7 signalling pathway and high IFN- α production. It is worth noting that the induction of interleukin-12 in cDCs stimulated by CpG-A was not markedly affected by DOTAP (Supplementary Fig. S3d), suggesting that the endosomal CpG-A signalling critical for the MyD88–IRF-7 pathway does not interfere with other pathways, such as the IRF-5 and NF- κ B pathways^{17,23}. Indeed, NF- κ B activation by CpG-A–DOTAP occurred in these stimulated cells and was unaffected by the absence of IRF-7 (Supplementary Fig. S3e).

The merger between CpG-A–Cy5 and the MyD88–IRF-7 complex was also observed in RAW cells 90 min after CpG-A–Cy5–DOTAP delivery (Fig. 4a, see also Fig. 2c), and the induction of high levels of IFN- α/β mRNAs was observed (Fig. 4b). In RAW cells coexpressing MyD88–CFP and IRF-7–YFP, stimulation with CpG-A–DOTAP also resulted in nuclear translocation of IRF-7–YFP (Supplementary Fig. S4a). However, IFN- α secretion from RAW cells was not observed (data not shown) and low levels of IFN- α were detected by intracellular staining with an anti-IFN- α antibody (Supplementary Fig. S4b), suggesting the lack of an additional post-transcriptional mechanism(s) in this cell line. Whatever the mechanism, low basal expression level of IRF-7 in these cells (data not shown) is sufficient to initiate and sustain MyD88–IRF-7 signalling required for robust IFN mRNA induction. Consistently, CD11b⁺B220[−] cells from bone marrow (myeloid lineage cells except pDCs) also secreted relatively high levels of IFN- α in response to CpG-A–DOTAP (Fig. 4c). Taken together, these results further support the endosomal trafficking model.

These observations prompted us to examine whether CpG-B, which *per se* cannot activate the MyD88–IRF-7 pathway, can also be manipulated to undergo endosomal trafficking and induce IFN- α by complexing with DOTAP. We found robust IFN- α induction in pDCs stimulated with CpG-B–DOTAP, which was also dependent

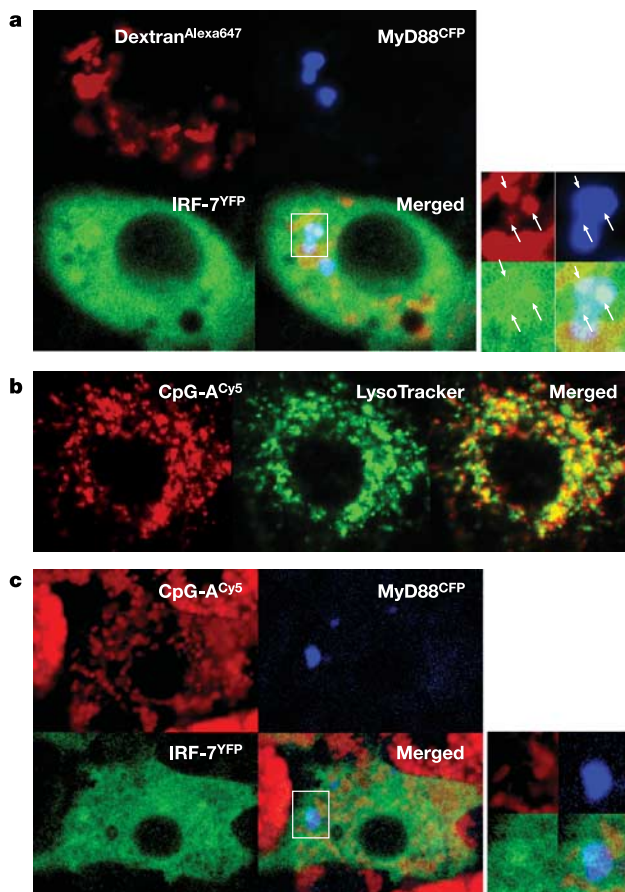


Figure 2 Subcellular localization of MyD88 and IRF-7, and trafficking of CpG ODNs in a macrophage cell line. **a**, RAW cells expressing MyD88–CFP and IRF-7–YFP were incubated with dextran–Alexa647 for 10 min. Higher magnification views (indicated by a white box) are shown in the bottom-right panel. White arrows indicate the colocalization of dextran–Alexa647, MyD88–CFP and IRF-7–YFP. **b**, Confocal images of RAW cells incubated with LysoTracker and 3 μ M CpG-A–Cy5 for 90 min. **c**, MyD88–CFP- and IRF-7–YFP-expressing RAW cells were incubated with 3 μ M CpG-A–Cy5 for 90 min. Note the absence of colocalization of CpG-A–Cy5 and MyD88–CFP/IRF-7–YFP.

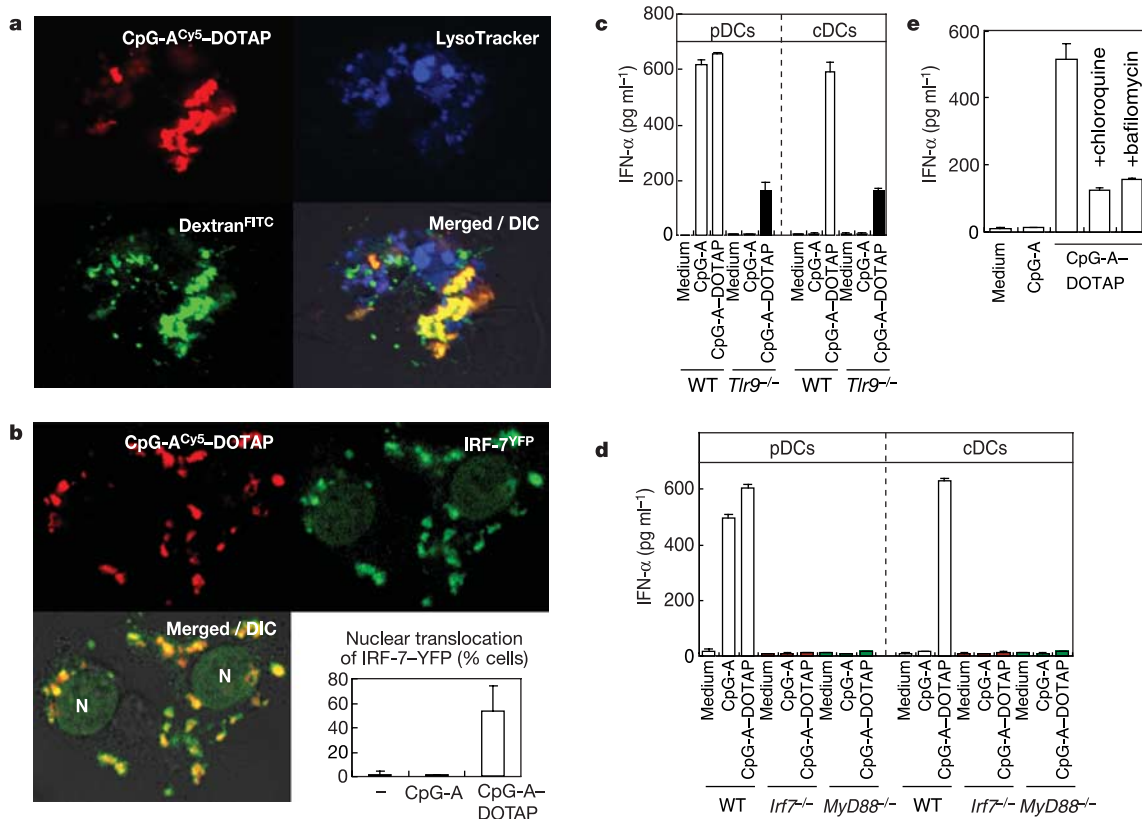


Figure 3 High production of type-I IFN in BM-derived cDCs achieved by manipulating CpG-A trafficking with DOTAP. **a**, cDCs were incubated with 3 μM CpG-A-Cy5-DOTAP for 80 min and, after washing, further incubated with LysoTracker and dextran-FITC for 10 min. **b**, cDCs expressing IRF-7-YFP were incubated with 3 μM CpG-A-Cy5-DOTAP for 90 min. N, nucleus. Cells showing the nuclear translocation of IRF-7-YFP were counted and the average of three independent experiments is shown (as mean ± s.d.,

right bottom panel). **c**, **d**, BM-derived DCs from wild-type (WT) or mutant mice were stimulated with 3 μM CpG-A or 3 μM CpG-A-DOTAP, or left untreated for 24 h as indicated. IFN-α induction was determined by ELISA. **e**, cDCs were stimulated with 3 μM CpG-A or 3 μM CpG-A-DOTAP in the presence or absence of inhibitors as indicated. Error bars in **c–e** represent the standard deviation of triplicate wells.

on TLR9, MyD88 and IRF-7 (Fig. 4d). This induction was dose-dependent, showing inhibition at high doses (Supplementary Fig. S4c and Supplementary Discussion). Furthermore, CpG-B-Cy5 now merged predominantly with dextran rather than with LysoTracker, but only at the optimal ligand concentration (Fig. 4e and Supplementary Fig. S4d). In contrast, CpG-B-DOTAP did not induce high IFN production in cDCs, even at the optimal concentration (Supplementary Fig. S4c). Consistent with the endosomal trafficking model, we observed that CpG-B-DOTAP merged predominantly with LysoTracker in cDCs (Supplementary Fig. S4e).

On the basis of our experimental data, we formulated a series of differential equations to explain the phenomena in this study (Supplementary Fig. S5). The kinetic model offers an explanation of how pDCs can achieve robust IFN induction when, for example, the difference in basal expression level of IRF-7 does not account for the induction (Supplementary Fig. S5c). The alteration of any parameters in the present model (such as TLR9 and MyD88) could not shift the equilibrium of CpG ODNs between endosomes and lysosomes (Supplementary Fig. S5d). Therefore, we propose the presence of an additional factor(s), which is specifically expressed in pDCs to modulate the equilibrium and retain the TLR9-bound CpG-A or CpG-B-DOTAP in endosomes.

The detailed mechanism by which DOTAP brings about the spatial merger with and the activation of MyD88-IRF-7 signalling by CpG ODNs is currently unknown. Cationic lipids can form a multilamellar structure with DNA and, under some conditions,

tend to form a coalescence with endosomal vesicles and remain stable^{18,24,25}. In this way, DOTAP might affect both the DNA structure and kinetics of DNA longevity in endosomes. As human peripheral blood mononuclear cells also produce high levels of IFN-α when stimulated by ssRNA complexed with DOTAP¹⁹, we infer from our results that DOTAP also influences the spatial control of ssRNA trafficking for activation of the TLR7-MyD88-IRF-7 pathway in endosomes^{20,21}. In this context, it is interesting to recall that nucleic acid/cationic lipid complexes may mimic viruses²⁵, and that viruses that activate the TLR9 subfamily accumulate in endosomes^{26,27}.

Our results might also have clinical implications. For example, in patients with systemic lupus erythematosus (SLE), endogenous IFN-α/β inducers consisting of the anti-DNA antibody/DNA complex are present, and these immune complexes bind to and enter into the endosomes of pDCs through Fc receptor-mediated endocytosis^{5,6,28}. In the light of our present study, we infer that, similar to DOTAP, the formation of complexes between DNA and its antibodies might mediate the endosomal retention of DNA, resulting in aberrant activation of the MyD88-IRF-7 pathway for IFN induction. It will be important to further study the nature of endosomal vesicles and their trafficking, in view of previous reports that indicate the highly heterogeneous nature of the vesicles²⁹.

The diversification of TLR9-MyD88 signalling to activate distinct transcription factors is critical to the orchestrated induction of distinct target genes that evoke adequate immune responses to

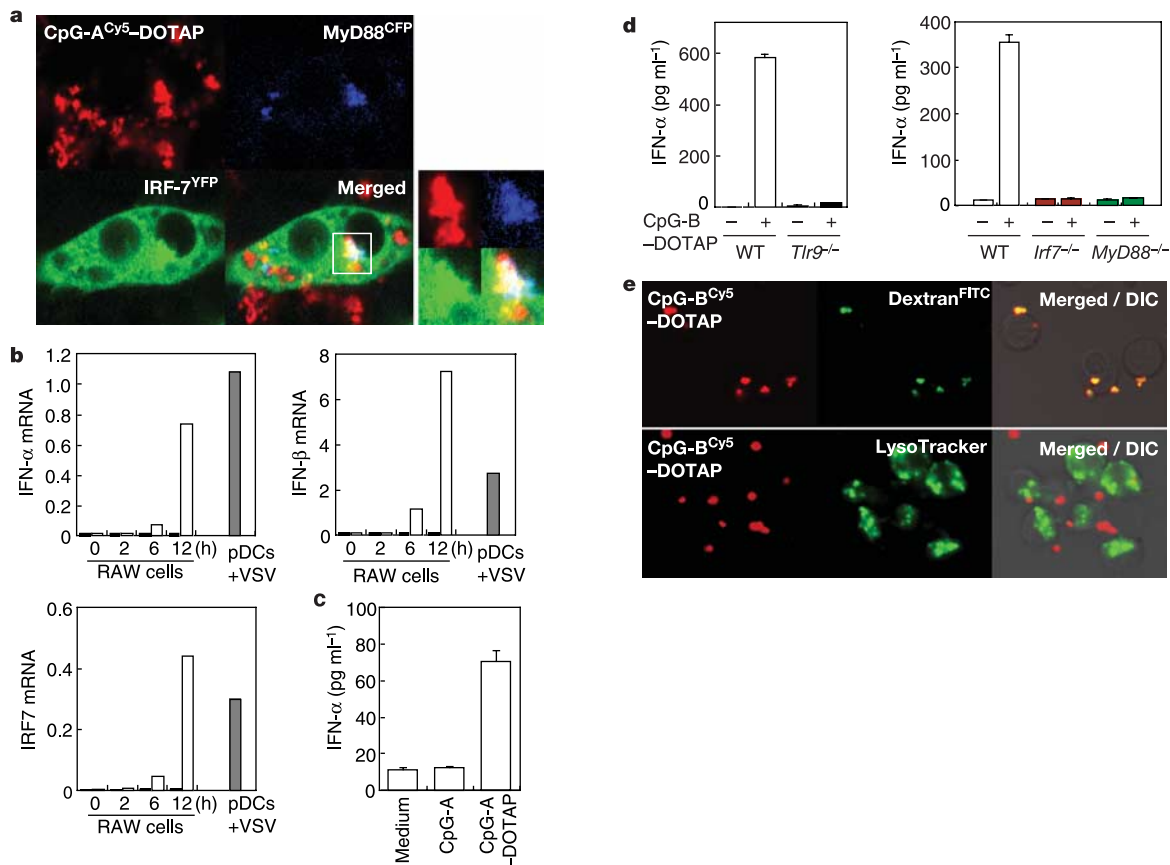


Figure 4 Induction of type-I IFN by manipulating ODN trafficking using DOTAP. **a**, MyD88-CFP- and IRF-7-YFP-expressing RAW cells were incubated with 3 μM CpG-A-Cy5-DOTAP for 90 min. **b**, RAW cells were stimulated with 3 μM CpG-A (black) or 3 μM CpG-A-DOTAP (white) for the indicated period. IFN-α/β and IRF-7 mRNA expression was monitored by quantitative real-time RT-PCR analysis. Data are normalized by the level of β-actin expression in each sample and shown as relative expression units. For comparison, splenic pDCs stimulated with vesicular stomatitis virus (VSV; 10 multiplicity of infection) for 9 h were also analysed. **c**, CD11b⁺B220⁺

cells purified directly from BM were stimulated with 3 μM CpG-A or 3 μM CpG-A-DOTAP. IFN-α levels were determined by ELISA. **d**, BM-derived pDCs from WT or mutant mice were stimulated with 1 μM CpG-B-DOTAP or left untreated for 24 h as indicated. IFN-α production was determined by ELISA. Error bars in **c** and **d** represent the standard deviation of triplicate cells. **e**, BM-derived pDCs were incubated with 1 μM CpG-B-Cy5-DOTAP, followed by incubation with dextran-FITC or LysoTracker.

infectious agents²⁻⁴. The spatiotemporal regulation described here might offer an explanation of how different types of unmethylated DNA differentially activate gene induction programs by restricting signalling cascades in space. Finally, our study might provide a rationale for the manipulation of DC-based immune responses, for example, by searching for compounds that activate or suppress the functions of targeted signalling vesicles. □

Methods

Reagents

The sequences of CpG ODNs are shown in Supplementary Methods. Cy5- and FITC-labelled CpG ODNs were purchased from Sigma Genosys and Hokkaido System Science, respectively. The complexation of CpG ODNs with DOTAP (Roche Diagnostics) was performed according to the manufacturer's instructions (see Supplementary Methods). FITC- and Alexa Fluor 647-labelled dextran, and LysoTracker Blue were purchased from Molecular Probes and used at final concentrations of 100 μg ml⁻¹ and 1 μM, respectively. Chloroquine and bafilomycin were purchased from Sigma and used at final concentrations of 5 μg ml⁻¹ and 30 nM, respectively.

Mice

The generation of *Ir7^{-/-}* mice has been described previously⁷. *MyD88^{-/-}* and *Tlr9^{-/-}* mice were provided by S. Akira. C57BL/6 mice were purchased from CLEA.

Preparation of DCs

To prepare bone-marrow-derived DCs, bone marrow cells were cultured for 6 days with 100 ng ml⁻¹ human Flt3L (Fms-like tyrosine kinase 3 ligand, PeproTech) in RPMI 1640 medium. Collected cells were incubated with the pDC-specific rat monoclonal antibody

120G8 (ref. 30; provided by G. Trinchieri) and anti-rat IgG microbeads (Miltenyi Biotec) and separated into pDCs (positively selected cells) and cDCs (flow-through cells) using a magnetic-activated cell sorting (MACS) column. In some experiments, collected cells were incubated with anti-B220 and anti-CD11c antibodies (BD Biosciences), and B220⁺/CD11c⁺ cDCs and B220⁺/CD11c^{intermediate} pDCs were sorted using FACS Diva (BD Biosciences). To prepare splenic pDCs, spleens were digested with collagenase A (Roche Biochemicals) and pDCs were purified using a pDC isolation kit (Miltenyi Biotec) according to the manufacturer's instructions.

Measurement of IFN production

Cells were seeded into 96-well plates at 2×10^5 cells ml⁻¹ and stimulated with various reagents for 24 h. Enzyme-linked immunosorbent assays (ELISAs) and quantitative real-time polymerase chain reaction with reverse transcription (RT-PCR) analysis were performed as described previously⁷.

Gene transduction

Details of lentiviral-mediated gene transfer are described in Supplementary Methods.

Imaging

Cells were cultured on glass-bottom, 35-mm tissue culture dishes (MATSUNAMI GLASS). Confocal microscopy studies were performed with an oil immersion objective (× 60 Plan Apo, numerical aperture 1.4) and an Olympus FV-1000 confocal microscope. Dual- or triple-colour images were acquired using a sequential acquisition mode to avoid cross-excitation.

Received 18 February; accepted 14 March 2005; doi:10.1038/nature03547.

Published online 6 April 2005.

- Colonna, M., Trinchieri, G. & Liu, Y. J. Plasmacytoid dendritic cells in immunity. *Nature Immunol.* **5**, 1219–1226 (2004).
- Wagner, H. The immunobiology of the TLR9 subfamily. *Trends Immunol.* **25**, 381–386 (2004).

3. Klinman, D. M. Immunotherapeutic uses of CpG oligodeoxynucleotides. *Nature Rev. Immunol.* **4**, 249–258 (2004).
4. Krieg, A. M. CpG motifs in bacterial DNA and their immune effects. *Annu. Rev. Immunol.* **20**, 709–760 (2002).
5. Theofilopoulos, A. N., Baccala, R., Beutler, B. & Kono, D. H. Type I interferons (α/β) in immunity and autoimmunity. *Annu. Rev. Immunol.* **23**, 307–335 (2005).
6. Ronnblom, L. & Alm, G. V. Systemic lupus erythematosus and the type I interferon system. *Arthritis Res. Ther.* **5**, 68–75 (2003).
7. Honda, K. *et al.* IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature* advance online publication, 30 March 2005 (doi:10.1038/nature03464).
8. Verthelyi, D., Ishii, K. J., Gursel, M., Takeshita, F. & Klinman, D. M. Human peripheral blood cells differentially recognize and respond to two distinct CpG motifs. *J. Immunol.* **166**, 2372–2377 (2001).
9. Krug, A. *et al.* Identification of CpG oligonucleotide sequences with high induction of IFN- α/β in plasmacytoid dendritic cells. *Eur. J. Immunol.* **31**, 2154–2163 (2001).
10. Hemmi, H., Kaisho, T., Takeda, K. & Akira, S. The roles of Toll-like receptor 9, MyD88, and DNA-dependent protein kinase catalytic subunit in the effects of two distinct CpG DNAs on dendritic cell subsets. *J. Immunol.* **170**, 3059–3064 (2003).
11. Kerkmann, M. *et al.* Activation with CpG-A and CpG-B oligonucleotides reveals two distinct regulatory pathways of type I IFN synthesis in human plasmacytoid dendritic cells. *J. Immunol.* **170**, 4465–4474 (2003).
12. Krieg, A. M. *et al.* CpG motifs in bacterial DNA trigger direct B-cell activation. *Nature* **374**, 546–549 (1995).
13. Izaguirre, A. *et al.* Comparative analysis of IRF and IFN- α expression in human plasmacytoid and monocyte-derived dendritic cells. *J. Leukoc. Biol.* **74**, 1125–1138 (2003).
14. Coccia, E. M. *et al.* Viral infection and Toll-like receptor agonists induce a differential expression of type I and λ interferons in human plasmacytoid and monocyte-derived dendritic cells. *Eur. J. Immunol.* **34**, 796–805 (2004).
15. Latz, E. *et al.* TLR9 signals after translocating from the ER to CpG DNA in the lysosome. *Nature Immunol.* **5**, 190–198 (2004).
16. Honda, K. *et al.* Role of a transductional-transcriptional processor complex involving MyD88 and IRF-7 in Toll-like receptor signaling. *Proc. Natl Acad. Sci. USA* **101**, 15416–15421 (2004).
17. Takaoka, A. *et al.* Integral role of IRF-5 in the gene induction programme activated by Toll-like receptors. *Nature* **434**, 243–249 (2005).
18. Zabner, J., Fasbender, A. J., Moninger, T., Poellinger, K. A. & Welsh, M. J. Cellular and molecular barriers to gene transfer by a cationic lipid. *J. Biol. Chem.* **270**, 18997–19007 (1995).
19. Heil, F. *et al.* Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science* **303**, 1526–1529 (2004).
20. Diebold, S. S., Kaisho, T., Hemmi, H., Akira, S. & Reis e Sousa, C. Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* **303**, 1529–1531 (2004).
21. Heil, F. *et al.* The Toll-like receptor 7 (TLR7)-specific stimulus loxoribine uncovers a strong relationship within the TLR7, 8 and 9 subfamily. *Eur. J. Immunol.* **33**, 2987–2997 (2003).
22. Hacker, H. *et al.* CpG-DNA-specific activation of antigen-presenting cells requires stress kinase activity and is preceded by non-specific endocytosis and endosomal maturation. *EMBO J.* **17**, 6230–6240 (1998).
23. Hacker, H. *et al.* Immune cell activation by bacterial CpG-DNA through myeloid differentiation marker 88 and tumor necrosis factor receptor-associated factor (TRAF)6. *J. Exp. Med.* **192**, 595–600 (2000).
24. Radler, J. O., Koltover, I., Salditt, T. & Safinya, C. R. Structure of DNA-cationic liposome complexes: DNA intercalation in multilamellar membranes in distinct interhelical packing regimes. *Science* **275**, 810–814 (1997).
25. Koltover, I., Salditt, T., Radler, J. O. & Safinya, C. R. An inverted hexagonal phase of cationic liposome–DNA complexes related to DNA release and delivery. *Science* **281**, 78–81 (1998).
26. Durrer, P., Gaudin, Y., Ruigrok, R. W., Graf, R. & Brunner, J. Photolabeling identifies a putative fusion domain in the envelope glycoprotein of rabies and vesicular stomatitis viruses. *J. Biol. Chem.* **270**, 17575–17581 (1995).
27. Brunetti, C. R., Dingwell, K. S., Wale, C., Graham, F. L. & Johnson, D. C. Herpes simplex virus gD and virions accumulate in endosomes by mannose 6-phosphate-dependent and -independent mechanisms. *J. Virol.* **72**, 3330–3339 (1998).
28. Leadbetter, E. A. *et al.* Chromatin–IgG complexes activate B cells by dual engagement of IgM and Toll-like receptors. *Nature* **416**, 603–607 (2002).
29. Mellman, I. Endocytosis and molecular sorting. *Annu. Rev. Cell Dev. Biol.* **12**, 575–625 (1996).
30. Asselin-Paturel, C., Brizard, G., Pin, J. J., Briere, F. & Trinchieri, G. Mouse strain differences in plasmacytoid dendritic cell frequency and function revealed by a novel monoclonal antibody. *J. Immunol.* **171**, 6466–6477 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank J. Vilcek, H. Rosen, H. Ohno, F. Nakatsu, A. Nakano, M. Lamphier, L. Cantley and T. Curran for advice, S. Akira for TLR9 and MyD88 mutant mice, G. Trinchieri for the pDC-specific antibody, A. Miyawaki for Venus (a variant of YFP), H. Miyoshi for lentivirus vectors, and M. Shishido for technical assistance. This work was supported in part by a grant for Advanced Research on Cancer and a Grant-In-Aid for Scientific Research on Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology of Japan, the Uehara Memorial Foundation, the Sumitomo Foundation, the Senri Life Science Foundation and the Nakajima Foundation. H.N. was supported by an Ishidu Shun Memorial Scholarship.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.T. (tada@m.u-tokyo.ac.jp).

Visualizing the mechanical activation of Src

Yingxiao Wang¹, Elliot L. Botvinick^{1,5}, Yihua Zhao¹, Michael W. Berns^{1,4,5}, Shunichi Usami¹, Roger Y. Tsien³ & Shu Chien^{1,2}

¹Department of Bioengineering and the Whitaker Institute of Biomedical Engineering, ²Department of Medicine, ³Howard Hughes Medical Institute and Departments of Pharmacology and Chemistry and Biochemistry, University of California at San Diego, La Jolla, California 92093, USA

⁴Department of Biomedical Engineering, ⁵Beckman Laser Institute, University of California at Irvine, Irvine, California 92697, USA

The mechanical environment crucially influences many cell functions¹. However, it remains largely mysterious how mechanical stimuli are transmitted into biochemical signals. Src is known to regulate the integrin–cytoskeleton interaction², which is essential for the transduction of mechanical stimuli^{3–5}. Using fluorescent resonance energy transfer (FRET), here we develop a genetically encoded Src reporter that enables the imaging and quantification of spatio-temporal activation of Src in live cells. We introduced a local mechanical stimulation to human umbilical vein endothelial cells (HUVECs) by applying laser-tweezer traction on fibronectin-coated beads adhering to the cells. Using the Src reporter, we observed a rapid distal Src activation and a slower directional wave propagation of Src activation along the plasma membrane. This wave propagated away from the stimulation site with a speed (mean $\pm$ s.e.m.) of 18.1 ± 1.7 nm s^{−1}. This force-induced directional and long-range activation of Src was abolished by the disruption of actin filaments or microtubules. Our reporter has thus made it possible to monitor mechanotransduction in live cells with spatio-temporal characterization. We find that the transmission of mechanically induced Src activation is a dynamic process that directs signals via the cytoskeleton to spatial destinations.

Mechanical stimuli activate integrins and the cytoskeleton to regulate cellular functions such as movement and adhesion⁶. When activated, integrins associate with Src via its SH3 domain, thus unmasking the Src kinase domain and activating Src⁷. Src can regulate integrin–cytoskeleton interaction², and cause dissolution of actin stress fibres and the release of mechanical tensile stress^{8,9}.

We previously developed a FRET-based Src indicator¹⁰, but that reporter also responded to Abl, Lck and epidermal growth factor receptor (EGFR). Furthermore, mutation of the Tyr within the substrate sequence to Phe did not eliminate the FRET response, which cast doubt on the mechanism for the FRET response. We therefore sought an indicator with higher specificity and a better-defined response mechanism. A Src substrate peptide (WMEDYDYVHLQG, derived from a primary *in vivo* c-Src substrate-molecule p130cas^{11,12}) was designed to replace the original substrate (EIYGEF, identified by *in vitro* library screening¹³) and to provide sufficient space for Src to gain access¹⁴ (Fig. 1a). The proximity of the N and C terminals of the SH2 domain, revealed by its crystal structure¹⁵, should allow the juxtaposition of cyan and yellow fluorescent proteins (CFP and YFP) to yield a high FRET. On Src phosphorylation, the substrate peptide can bind to the phosphopeptide-binding pocket of the SH2 domain and separate YFP from CFP, thus decreasing the FRET (Fig. 1b). Phosphorylation of the purified reporter by Src *in vitro* enhanced CFP emission at the expense of YFP emission (Fig. 1c) and increased the cyan-to-yellow emission ratio by 25%, indicating a Src-induced loss of FRET. The specificity was very much better than for the previous Src indicator. The emission ratio changed by <2% for other kinases (Yes, FAK, EGFR, Abl, Jak2 or Ser/Thr kinase ERK1) and changed moderately

(~10%) only for Fyn, a close relative of Src (Fig. 1d). The Src-induced loss of FRET is consistent with intramolecular complexation of the phosphorylated substrate with the SH2 domain and the consequent disruption of the close apposition of the CFP and YFP domains (Fig. 1b).

In HeLa cells transfected with the Src reporter, epidermal growth factor (EGF) induced a 25–35% emission ratio change (Fig. 2a; Supplementary Movie 1). Introduction of the reporter did not affect the ERK activity with or without EGF stimulation (data not shown), suggesting that the reporter need not perturb endogenous cellular signalling. Mutations of either or both of the putative Src phosphorylation sites (Tyr 662 and 664) to Phe in the substrate peptide (Fig. 1a) prevented the FRET response to EGF in HeLa cells (Fig. 2b). Mutation of Arg 175 to Val (Fig. 1a), eliminating SH2 domain binding to phosphorylated peptides¹⁵, also abrogated the EGF-induced FRET response. These results validate the phosphorylation-induced intramolecular (intra-reporter) interaction between the SH2 domain and substrate peptide as the mechanism for the FRET response. Immunoblotting revealed that EGF-induced tyrosine phosphorylation is abolished only by mutating both Tyr 662 and 664, but not either site alone (Fig. 2c), unlike the blockade of the FRET response by single mutations (Fig. 2b). Thus, the SH2 binding requires not only the phosphorylation of one of the two Tyr residues but also the integrity of the other Tyr in the substrate, consistent with the notion that the neighbouring amino acids of the phosphorylated site are important for SH2 binding¹⁶. Disruption of the SH2 domain by R175V mutation also blocked the EGF-induced tyrosine phosphorylation of the Src reporter (Fig. 2c), suggesting that the SH2 domain of the reporter may assist in its association with activated Src to facilitate the phosphorylation process.

The EGF-induced FRET response in HeLa cells was reversed by PP1, a selective inhibitor of Src family tyrosine kinases, and was markedly reduced after pretreatment with PP1 (Fig. 2d). The

normal platelet-derived growth factor (PDGF)-induced FRET response of the Src reporter was abolished in Src/Yes/Fyn triple-knockout (SYF^{-/-}) mouse embryonic fibroblasts (MEF) (Fig. 2e). Reconstitution of these SYF^{-/-} cells with c-Src, but not c-Fyn, restored the FRET response. Reconstitution with c-Yes caused only a weak FRET response of the Src reporter. The small and delayed FRET response observed in the kinase-dead c-Src (Src_KD) group may be attributed to the residual activity of Src_KD. These results demonstrated the specificity of the Src reporter toward Src in mammalian cells.

CFP and YFP can form anti-parallel dimers¹⁷. To eliminate the unintended FRET resulting from intermolecular (between reporters) dimerization, we introduced A206K mutations into CFP and YFP to generate monomeric CFP and YFP¹⁸ (Fig. 1a). These mutations did not alter the spectral properties of the Src reporter, but they led to a better dynamic range of FRET (43 versus 25% emission ratio change) in response to Src kinase *in vitro* (compare Fig. 2f with Fig. 1c) and a greater EGF-induced FRET response in HeLa cells that was reversible by EGF washout (Fig. 2g, Supplementary Fig. 1 and Supplementary Movie 2). This monomeric reporter was used for the study of mechano-activation of Src.

Beads coated with fibronectin, which binds to integrins and hence causes coupling with the cytoskeleton¹⁹, were applied to HUVECs. Consistent with the observation of Src activation by integrin clustering⁷, the fibronectin-coated beads caused a local FRET response of the Src reporter around the beads (Fig. 3a and Supplementary Movie 3), reversible by PP1 (Supplementary Fig. 2 and Supplementary Movie 4). Single-beam gradient optical laser-tweezers with controlled mechanical force (300 pN) were used to pull the adhered beads. FRET responses occurred in focal-complex-like regions at the cell periphery without detectable bead displacement (Supplementary Fig. 3a and Supplementary Movie 5). Polylysine-coated beads subjected to the same mechanical force did

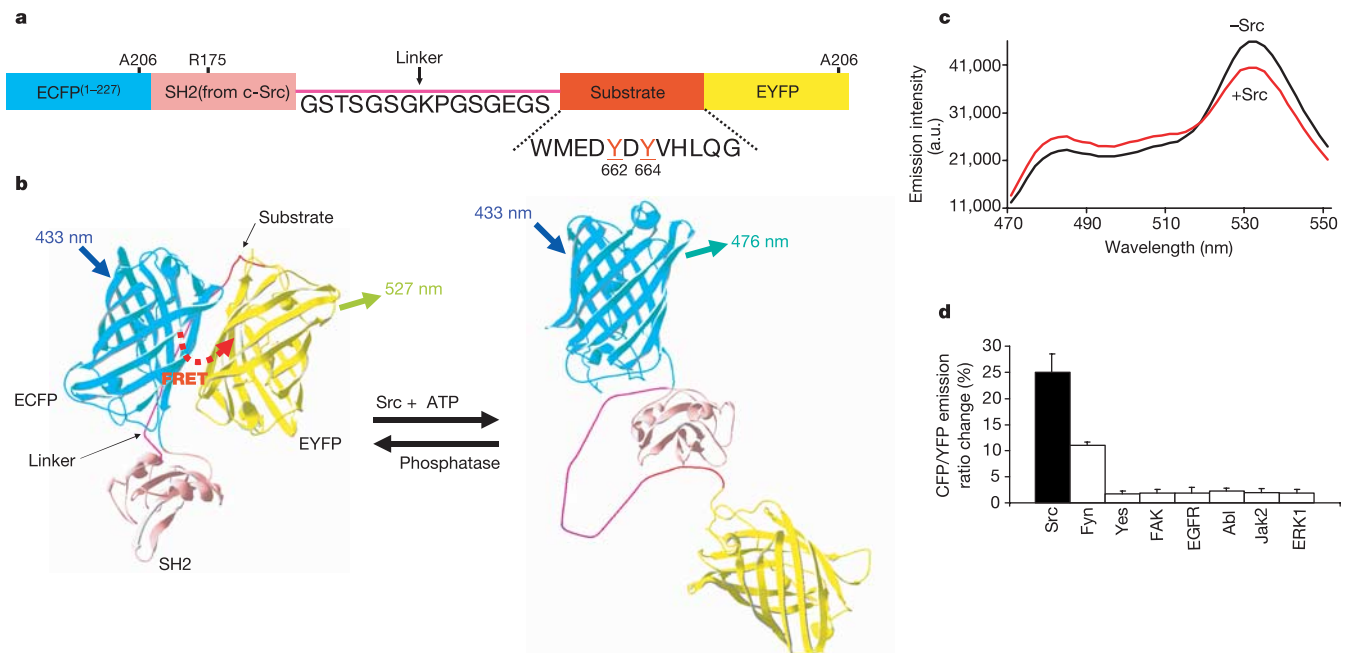


Figure 1 The domain structure, schematic representation and *in vitro* characterizations of the Src reporter. **a**, The Src reporter is composed of CFP, the SH2 domain, a flexible linker, the Src substrate peptide and YFP (ref. 10). **b**, The cartoon illustrates the FRET effect of the Src reporter upon the actions of Src kinase or phosphatase. **c**, Emission

spectra of the Src reporter before (black) and after (red) phosphorylation by Src. **d**, *In vitro* emission ratio changes (mean $\pm$ s.d.) of the Src reporter in response to Src and other kinases.

not induce any significant FRET response (Supplementary Fig. 3b and Supplementary Movie 6), suggesting that specific integrin–cytoskeleton coupling is needed for the mechanotransduction.

The thin lamellipodia at the cell periphery, which are important in mechanotransduction, contain only limited copies of cytosolic Src reporters. Because significant amounts of CFP/YFP molecules are required to yield enough fluorescence above the endogenous autofluorescence background²⁰, there is a need for controlled localization of the Src reporter to enhance its effective local concentration, especially in lamellipodia. Because the translocation of Src to the plasma membrane is a prerequisite for Src activation^{7,21}, we targeted the monomeric Src reporter to the plasma membrane with a fusion of the 16 N-terminal residues from Lyn kinase¹⁸. The EGF-induced FRET response of this membrane-targeted reporter was reversed by PP1 (Supplementary Fig. 4a and Supplementary Movie 7) and prevented by pretreatment with PP1 (Supplementary Fig. 4b and Supplementary Movie 8), indicating its specificity towards Src (Fig. 3b).

The application of pulling force via the laser tweezers on a bead coated with fibronectin, but not polylysine (Supplementary Fig. 5), on the HUVECs expressing the membrane-targeted Src reporter led to a directional FRET response, with the majority of activations

transmitted towards distal areas of the cell opposite to the force direction (Fig. 3c; Supplementary Movie 9). This transmission consisted of an immediate distal Src activation and a slower wave-propagation of Src activation followed by lamellipodia protrusions at the cell periphery (Fig. 3c). Such FRET responses were absent with inactive reporters (for Y662F/Y664F see Supplementary Fig. 6a and Supplementary Movie 10; for R175V see Supplementary Fig. 6b and Supplementary Movie 11). In experiments where the speed of the wave propagation of Src-activation away from the stimulation site can be clearly measured, it was found to be $18.1 \pm 1.7 \text{ nm s}^{-1}$ (mean $\pm$ s.e.m.) (Fig. 3d; Supplementary movie 12). This result indicates that a local force caused a directional and long-range transduction of Src-activation wave to spatial destinations.

We examined the roles of the cytoskeleton on this force-induced Src activation. Disruption of actin filaments with cytochalasin D or microtubules with nocodazole blocked the force-induced distal, but not local, FRET responses (Fig. 4a, b; Supplementary Movies 13, 14). Polarity analysis of Src activation, by averaging the emission ratios of 36 evenly divided angular sections of each cell with the bead position as the centre, revealed that the local pulling force caused a cytoskeleton-dependent polarized

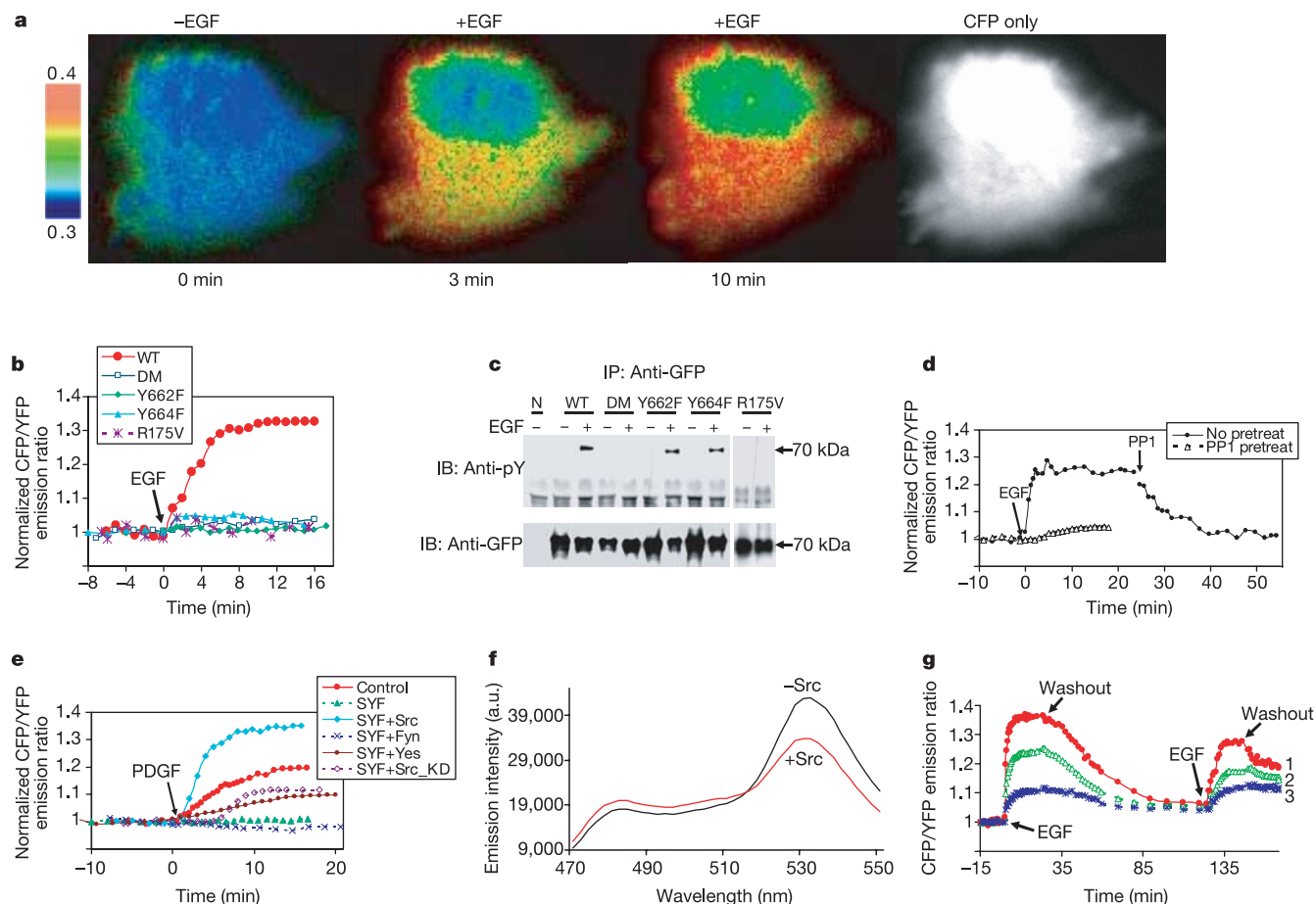


Figure 2 Characterization of the Src reporters. **a**, The CFP/YFP emission ratio images in response to EGF (Supplementary Movie 1). **b**, Emission ratio time courses of the Src reporter and its mutants (see Methods) in response to EGF stimulation in HeLa cells. **c**, The tyrosine phosphorylation level of the various Src reporters (see Methods). 'N' represents cells without transfection. **d**, Emission ratio time courses of the Src reporter in response to EGF in HeLa cells pretreated with ('PP1 pretreat') or without ('No pretreat')

PP1. **e**, Emission ratio time courses of the Src reporter in response to PDGF in MEF and SYF cell lines (see Methods). **f**, Emission spectra of the monomeric Src reporter before (black) and after (red) *in vitro* phosphorylation by Src. **g**, Emission ratios of the monomeric Src reporter from three different subcellular regions in HeLa cells stimulated by EGF, followed sequentially by EGF washout, re-stimulation, and a second washout (Supplementary Movie 2).

FRET response pointing to the opposite direction (Fig. 4c, d). Statistical analysis further showed that the mechanical force caused a cytoskeleton-dependent increase of the number of pixels with highly activated Src distal to the force-imposed bead, indicating a long-range mechano-activation of Src (Fig. 4e; Supplementary Fig. 7).

It is unclear where mechano-induced biochemical signals are initiated and how they are transmitted in the cell. Green fluorescent protein (GFP)-tagged fluorescent markers have been used to study the displacement of cellular organelles and the formation of a focal adhesion complex induced by mechanical stimuli^{22–25}, but these inert fluorescence markers cannot monitor the dynamic signal transduction process. Our FRET-based Src reporter enables the visualization and quantification of the mechano-activated Src with high temporal and spatial resolution in live cells. The results indicate that local mechanical stimulation triggers a directional and

long-range propagation of Src activation, for which cytoskeleton integrity is essential.

Integrin-mediated activation of Src at local sites by mechanical stimuli may induce p130cas/Dock180 association, Rac-Arp2/3 activation, cortical actin network nucleation and polymerization, and actin-ruffle extensions^{26,27}. These Rac and actin activities in turn promote the recruitment and activation of Src at the tip of these newly assembled wave-like actin ruffles^{28,29}, thus further inducing the *in situ* Rac activation and actin polymerization. This positively coordinated mechanism may result in a wave propagation of Src activation. The directionality of this wave propagation may be attributed to the initial local mechanical tension generated in a direction counter to the applied force. The applied force can also be mechanically transmitted quickly through tensed cytoskeleton network to distal locations and to activate Src¹. This directional Src activation may release the tension²¹ at desired destinations and

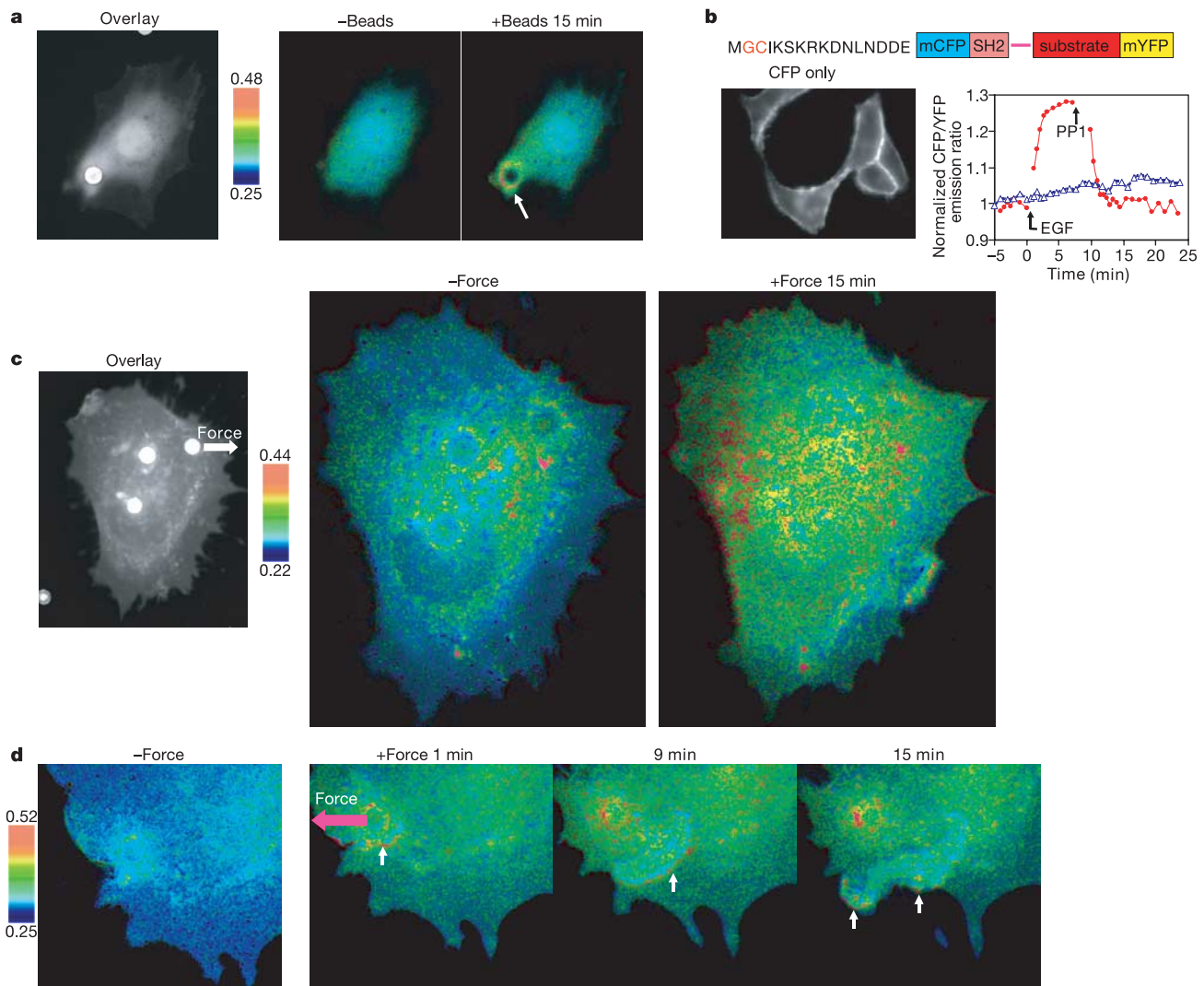


Figure 3 Directional and long-range propagation of Src induced by mechanical force. **a**, A fibronectin-coated bead (white spot from phase contrast image overlaid on CFP cell image) induced FRET responses around the bead (Supplementary Movie 3). White arrow points to the spot with activated Src. Colour bar represents CFP/YFP emission ratio values. **b**, The schematic diagram in the upper panel shows the design strategy of membrane targeting. The CFP-only image on the left shows the effective tethering of the reporter on the plasma membrane. The EGF-induced FRET responses of the reporter is reversed by

PP1 (red line; Supplementary Movie 7) and prevented by pretreatment with PP1 (blue line; Supplementary Movie 8). **c**, Laser tweezer traction on the bead at the upper right corner of the cell (shown on the left) caused FRET responses (Supplementary Movie 9). White arrow represents force direction. **d**, FRET responses of a cell with clear directional wave propagation away from the site of mechanical stimulation (Supplementary Movie 12).

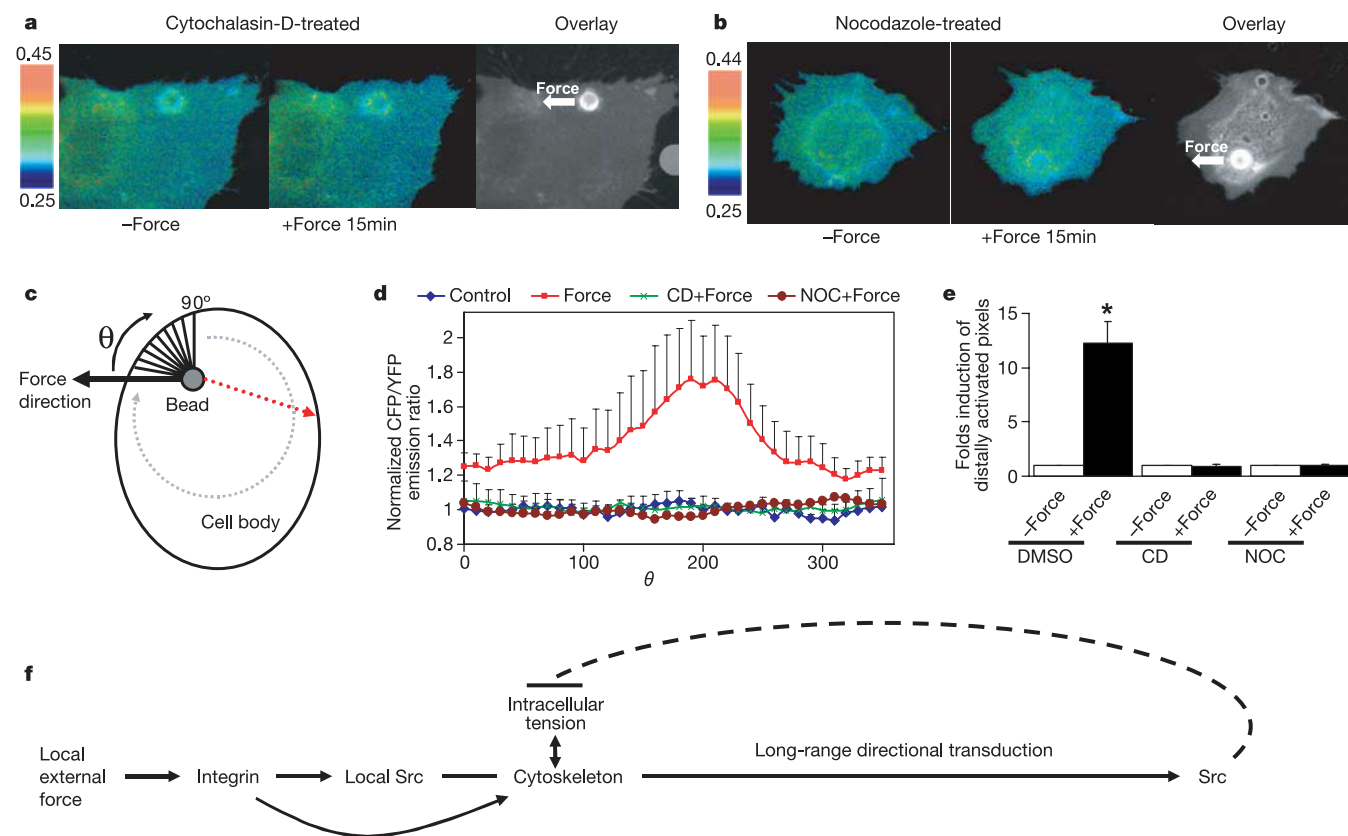


Figure 4 Actin filaments and microtubules are essential for the polarized and long-range Src activation. Emission ratio images of HUVECs pretreated with cytochalasin D (**a**, Supplementary Movie 13) or nocodazole (**b**, Supplementary Movie 14) before and after force application. **c**, Schematic drawing of the polarity analysis strategy. **d**, Polarity analysis (mean $\pm$ s.d.) of the force-induced FRET response. Control, no force application. CD or NOC, cytochalasin D or nocodazole treatment, respectively. **e**, HUVECs treated with

cytochalasin D, nocodazole or DMSO were subjected to mechanical force for 15 min or kept as static control. Bar graphs represent mean $\pm$ s.d. of the force-induced fold induction of distally activated (see Methods, with 80% threshold) pixel numbers. The asterisk indicates a significant difference ($P < 0.05$) before and after force application. **f**, A proposed model depicting the mechanism by which local mechanical forces induce directional and long-range Src activations.

rearrange the intracellular stress distribution, thus serving as a feedback mechanism for the cell to adapt to new mechanical environments (Fig. 4f).

Methods

Laser-tweezers

A 1064-nm continuous-wave diode-pumped ND:YV0₄ laser with 5W power (Spectra-Physics) was used for the laser-tweezers experiments. The laser beam passes through a laser-beam expander, a steering mirror, and a dichroic long-pass beamsplitter to enter the microscope side port.

Gene construction and DNA plasmids

The gene for the Src reporter was constructed by polymerase chain reaction (PCR) amplification of the complementary DNA from the c-Src SH2 domain with a sense primer containing a *SphI* site and a reverse primer containing the gene sequence for a flexible linker, a substrate peptide derived from p130cas, and a *SacI* site. The PCR products were fused together with an N-terminal enhanced CFP and a C-terminal citrine (a version of enhanced YFP)¹⁰, as shown in Fig. 1a. Mutations of Y662/664F, Y662F, Y664F, R175V and A206K were conducted with the QuickChange method (Stratagene). Constructs were cloned into pRSET_B (Invitrogen) using *Bam*HI/*Eco*RI for bacterial expression and into pcDNA3 (Invitrogen) behind a Kozak sequence using *Hind*III/*Eco*RI for mammalian cell expression. The membrane-targeted CFP was constructed by PCR amplification of the monomeric CFP with a sense primer containing the codes for 16 N-terminal amino acids from Lyn kinase¹⁸ to produce a membrane-targeted Src reporter.

The various Src reporters and their mutants used in Figs 2b and c are abbreviated as: WT, the Src reporter (wild type); DM, Y662F and Y664F double mutations in the designed substrate peptide; Y662F or Y664F, the Y662F or Y664F single mutation in the substrate peptide, respectively; and R175V, the R175V mutation in the binding pocket of the SH2 domain.

Cell lines

The various mouse embryonic fibroblasts (MEFs) and knockout cell lines used for specificity studies in Fig. 2e are: wild type (control), Src/Yes/Fyn triple-knockout (SYF), SYF reconstituted with c-Src (SYF+Src), c-Fyn (SYF+Fyn), c-Yes (SYF+Yes), or K295R kinase-dead c-Src (SYF+Src_KD).

Microscopy and image acquisition

The HeLa or MEF cells expressing the desired exogenous proteins were starved with 0.5% FBS for 36–48 h before being subjected to EGF (50 ng ml⁻¹) or PDGF (10 ng ml⁻¹) stimulation. During imaging, the cells were maintained in Hanks' balanced salt solution (HBSS) with 20 mM HEPES (pH 7.4) and 2 g l⁻¹ D-glucose at 25 °C. Images were collected by using MetaFluor 6.0 software (Universal Imaging) with a 440DF20 excitation filter, a 455DRLP dichroic mirror, and two emission filters controlled by a filter changer (480DF30 for CFP and 535DF25 for YFP).

To image the mechanical-force-induced Src activation, HUVECs were first starved with 0.5% FBS for 24 h and then kept in CO₂-independent medium without serum (Gibco BRL) at 37 °C in a thermostatic chamber. A Zeiss axiovert inverted microscope equipped with a 440DF20 excitation filter and a 455DRLP dichroic mirror was integrated with the laser-tweezers. CFP and YFP emission images were acquired simultaneously with an ORCA ER CCD camera (Hamamatsu) through a Dual-View module (Optical-Insights). The CFP and YFP images were aligned pixel-by-pixel with our customized Matlab program by maximizing the normalized cross-correlation coefficient of CFP and YFP fluorescence intensity images:

$$\text{corr} = \frac{\sum_i \sum_j (C_{ij} - \bar{C})(Y_{ij} - \bar{Y})}{\sqrt{\left(\sum_i \sum_j (C_{ij} - \bar{C})^2 \right) \left(\sum_i \sum_j (Y_{ij} - \bar{Y})^2 \right)}}$$

Where C_{ij} and Y_{ij} are the intensity values at pixels (i, j) of the CFP and YFP images, and $\bar{C}$ and $\bar{Y}$ are the mean intensity values of the CFP and YFP images. The ratio images of aligned CFP/YFP were computed and created by the MetaFluor software to represent the FRET efficiency.

Image analysis

By taking the derivative of $\sigma_{\theta}^2(\kappa)$ in Otsu's method³⁰ and calculating its first local minimum, a non-parametric method was developed to calculate the intensity threshold to differentiate the edges of HUVECs from the background in CFP or YFP fluorescence intensity images. The intensity threshold was used to generate a binary mask image with values outside the cell set at zero to select the pixels located within the cell body in the CFP/YFP emission ratio images. For polarity analysis, a customized Matlab program was used to evenly divide a HUVEC into 36 angular sections with the bead position as the centre and the force direction as the zero degree axes for θ , as illustrated in Fig. 4c. For statistical analysis of long-range activation of Src, a pixel located within a cell at a given time is defined as 'distally activated' when: (1) its emission ratio value is above a certain percentage (98%, 90%, 80% and 50% were used as the thresholds) of pixel population in the same cell after 15 min of force stimulation, and (2) the distance between the pixel and the centre of the force-imposed bead is larger than half of the virtual radius of the cell, as shown in the following formula:

$$\sqrt{(x-x_0)^2+(y-y_0)^2} > \frac{R}{2} = \frac{1}{2} \sqrt{\frac{A}{\pi}}$$

where x, y are the coordinates of any given pixel within the cell body, x_0, y_0 are the coordinates of the force-imposed bead centre, R is the virtual radius for the cell, and A is the area of the cell.

Received 2 November 2004; accepted 11 February 2005; doi:10.1038/nature03469.

- Ingber, D. E. Tensegrity. II. How structural networks influence cellular information processing networks. *J. Cell. Sci.* **116**, 1397–1408 (2003).
- Felsenfeld, D. P., Schwartzberg, P. L., Venegas, A., Tse, R. & Sheetz, M. P. Selective regulation of integrin–cytoskeleton interactions by the tyrosine kinase Src. *Nature Cell Biol.* **1**, 200–206 (1999).
- Critchley, D. R. Focal adhesions—the cytoskeletal connection. *Curr. Opin. Cell Biol.* **12**, 133–139 (2000).
- Chicurel, M. E., Singer, R. H., Meyer, C. J. & Ingber, D. E. Integrin binding and mechanical tension induce movement of mRNA and ribosomes to focal adhesions. *Nature* **392**, 730–733 (1998).
- Jiang, G., Giannone, G., Critchley, D. R., Fukumoto, E. & Sheetz, M. P. Two-piconewton slip bond between fibronectin and the cytoskeleton depends on talin. *Nature* **424**, 334–337 (2003).
- Alenghat, F. J. & Ingber, D. E. Mechanotransduction: all signals point to cytoskeleton, matrix, and integrins. *Sci. STKE* **2002**, E6 (2002).
- Arias-Salgado, E. G. *et al.* Src kinase activation by direct interaction with the integrin beta cytoplasmic domain. *Proc. Natl Acad. Sci. USA* **100**, 13298–13302 (2003).
- Fincham, V. J., Brunton, V. G. & Frame, M. C. The SH3 domain directs acto-myosin-dependent targeting of v-Src to focal adhesions via phosphatidylinositol 3-kinase. *Mol. Cell Biol.* **20**, 6518–6536 (2000).
- Kaplan, G. *et al.* Identification of a surface glycoprotein on African green monkey kidney cells as a receptor for hepatitis A virus. *EMBO J.* **15**, 4282–4296 (1996).
- Ting, A. Y., Kain, K. H., Klemke, R. L. & Tsien, R. Y. Genetically encoded fluorescent reporters of protein tyrosine kinase activities in living cells. *Proc. Natl Acad. Sci. USA* **98**, 15003–15008 (2001).
- Hamasaki, K. *et al.* Src kinase plays an essential role in integrin-mediated tyrosine phosphorylation of Crk-associated substrate p130Cas. *Biochem. Biophys. Res. Commun.* **222**, 338–343 (1996).
- Vuori, K., Hirai, H., Aizawa, S. & Ruoslahti, E. Introduction of p130cas signaling complex formation upon integrin-mediated cell adhesion: a role for Src family kinases. *Mol. Cell Biol.* **16**, 2606–2613 (1996).
- Songyang, Z. *et al.* Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature* **373**, 536–539 (1995).
- Yang, E., Liu, Y., Bixby, S. D., Friedman, J. D. & Shokat, K. M. Highly efficient green fluorescent protein-based kinase substrates. *Anal. Biochem.* **266**, 167–173 (1999).
- Waksman, G. *et al.* Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. *Nature* **358**, 646–653 (1992).
- Bradshaw, J. M. & Waksman, G. Calorimetric examination of high-affinity Src SH2 domain-tyrosyl phosphopeptide binding: dissection of the phosphopeptide sequence specificity and coupling energetics. *Biochemistry* **38**, 5147–5154 (1999).
- Yang, F., Moss, L. G. & Phillips, G. N. Jr The molecular structure of green fluorescent protein. *Nature Biotechnol.* **14**, 1246–1251 (1996).
- Zacharias, D. A., Violin, J. D., Newton, A. C. & Tsien, R. Y. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **296**, 913–916 (2002).
- Miyamoto, S. *et al.* Integrin function: molecular hierarchies of cytoskeletal and signaling molecules. *J. Cell Biol.* **131**, 791–805 (1995).
- Tsien, R. Y. The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509–544 (1998).
- Thomas, S. M. & Brugge, J. S. Cellular functions regulated by Src family kinases. *Annu. Rev. Cell Dev. Biol.* **13**, 513–609 (1997).
- Heidermann, S. R., Kaech, S., Buxbaum, R. E. & Matus, A. Direct observations of the mechanical behaviors of the cytoskeleton in living fibroblasts. *J. Cell Biol.* **145**, 109–122 (1999).
- Hu, S. *et al.* Intracellular stress tomography reveals stress focusing and structural anisotropy in cytoskeleton of living cells. *Am. J. Physiol. Cell Physiol.* **285**, C1082–C1090 (2003).
- Mack, P. J., Kaazempur-Mofrad, M. R., Karcher, H., Lee, R. T. & Kamm, R. D. Force-induced focal adhesion translocation: effects of force amplitude and frequency. *Am. J. Physiol. Cell Physiol.* **287**, C954–C962 (2004).
- Riveline, D. *et al.* Focal contacts as mechanosensors: externally applied local mechanical force induces growth of focal contacts by an mDia1-dependent and ROCK-independent mechanism. *J. Cell Biol.* **153**, 1175–1186 (2001).
- Brugnera, E. *et al.* Unconventional Rac-GEF activity is mediated through the Dock180-ELMO complex. *Nature Cell Biol.* **4**, 574–582 (2002).
- Miki, H., Suetsugu, S. & Takenawa, T. WAVE, a novel WASP-family protein involved in actin reorganization induced by Rac. *EMBO J.* **17**, 6932–6941 (1998).
- Sandilands, E. *et al.* RhoB and actin polymerization coordinate Src activation with endosome-mediated delivery to the membrane. *Dev. Cell* **7**, 855–869 (2004).
- Timpson, P., Jones, G. E., Frame, M. C. & Brunton, V. G. Coordination of cell polarization and migration by the Rho family GTPases requires Src tyrosine kinase activity. *Curr. Biol.* **11**, 1836–1846 (2001).
- Otsu, N. Threshold selection method from Gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **9**, 62–66 (1979).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank J. Zhang and A. Y. Ting for their assistance and advice; J.-L. Guan, D. C. Flynn and J. Cooper for providing FAK protein, MEF cells, c-Yes plasmid and SYF cell lines. We also thank B. N. G. Giepmans and C. Hauser for their comments on the paper, and Q. Xiong for technical support. This work was supported by the National Institutes of Health (S.C., R.Y.T. and M.W.B.), the Howard Hughes Medical Institute (R.Y.T.), the Airforce Office of Scientific Research (M.W.B.), the Arnold and Mabel Beckman Foundation (E.L.B.) and the Alliance for Cellular Signaling Grant (R.Y.T., S.C. and Y.W.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.C. (schien@bioeng.ucsd.edu). The sequence reported in this paper has been deposited in the GenBank database (accession number AY862145).

Contacts

Publisher: Ben Crowe

Editor: Paul Smaglik

Marketing Manager: David Bowen

US Head Office, New York

345 Park Avenue South, 10th Floor,

New York, NY 10010-1707

Tel +1 800 989 7718

Fax +1 800 989 7103

e-mail: naturejobs@natureny.com

US Sales Manager/ Corporations:

Peter Bless

Classified Sales Representatives:

Phone: +1 800 989 7718

New York/ Pennsylvania/

Latin America:

Kelly Roman

Midwest USA/ Maryland/ NIH:

Wade Tucker

East USA/ Canada:

Janine Taormina

San Francisco Office

Classified Sales Representative:

Michaela Bjorkman

West USA/ West Corp. Canada

225 Bush Street, Suite 1453

San Francisco, CA 94104

Tel +1 415 781 3803

Fax +1 415 781 3805

e-mail: m.bjorkman@nature.sf.com

European Head Office, London

The Macmillan Building, 4 Crinan Street,

London N1 9XW, UK

Tel +44 (0) 20 7843 4961

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Naturejobs Sales Director:

Nevin Bayoumi (4978)

European Sales Manager:

Andy Douglas (4975)

Advertising Production

Manager: Billie Franklin

To send materials use London

address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Niamh Shields

European Satellite Office

Germany/ Austria/ Italy/

The Netherlands/ Belgium:

Patrick Phelan

e-mail: p.phelan@nature.com

Reya Silao

e-mail: r.silao@nature.com

Japan Head Office, Tokyo

MG Ichigaya Building (5F),

19-1 Haraikatamachi,

Shinjuku-ku,

Tokyo 162-0841

Tel +81 3 3267 8751

Fax +81 3 3267 8746

Asia-Pacific Sales Director:

Rinoko Asami

e-mail: rasami@nature.jp

naturejobs

Breaking the bottleneck

It's not surprising to hear that postdocs are spending an increasing amount of their careers as fellows. Stories have circulated for years about people doing two or even three consecutive fellowships before going it alone. In 2002, for example, US researchers achieved their independence at a median age of 42, according to the National Institutes of Health. But a proposal, released last month by the US National Research Council (NRC), urges a multipronged approach to break the bottleneck between training and independence for young scientists.

The proposal puts forward some bold suggestions. First, it says that funding bodies should cap fellowships at five years. If a lab wants to keep the fellows for longer, it should reclassify them as 'staff scientists', and pay them accordingly.

Second, funding bodies should create more 'bridge' awards to help postdocs escape from the need for their senior investigator's primary grant. Some such awards exist, but many of them are confusing and need to be expanded to be more broadly applicable, says the NRC. There could be a new system, the council suggests, perhaps one that awards 200 investigators \$500,000 over five years.

Finally, these grants should be available to all, not just to young scientists who have preliminary data that predict a successful project. Instead, special new investigator grants should allow them to produce such data and give them a way out of this chicken-and-egg conundrum.

All these proposals offer sensible solutions to a problem that most agree exists — although two things are still missing: money and commitment. Money is limited because of the fickle nature of yearly budget cycles. But the commitment can and should come from senior scientists who acknowledge that a bottleneck exists and are determined to break it.

Paul Smaglik
Naturejobs editor



Contents

POSTDOCS & STUDENTS

Keeping control of your money p1048

CAREER VIEW

Scientists & Societies

Postdoc problems —

postdoc solution

Graduate Journal

First-night nerves

Movers

Tim Berners-Lee

p1050

WWW.NATUREJOBS.COM

Career centre

Information on the scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

ANNOUNCEMENTS

EVENTS

Counting the cost

Budgeting wisely can be critical to a young lab's early success. Kendall Powell finds out how to stretch funds for maximum effect.

Christina Hull knows that it takes money to make money. She's starting a small business with some venture capital which she must "convert into results before the investors pull the plug". But in this case, the small business is her molecular biology lab at the University of Wisconsin, Madison, and the venture money comes from her start-up package and a Burroughs Wellcome Career Award transition grant.

"It is a small business and it needs to be treated that way," Hull says of an academic laboratory. Researchers ought to think of research funding as being like a venture-capital financing package, says Hull, who took up her faculty position 20 months ago.

Michael McClure, semi-retired from a 23-year career at the National Institutes of Health, couldn't agree more. "Success is actually tied more to your business acumen than to your intellectual ability" in ultra-competitive research, says McClure, president of Frontiers Fund in Durham, North Carolina, which supports career development for young scientists.

Although many postdocs have written fellowship proposals, most have never set or spent their own research budget. The change triggered by a start-up or major grant budget can be overwhelming. Both new and established investigators say that, with a bit of forethought, postdocs making the transition can learn to get the biggest bang from initial research bucks. Leveraging sources of funding against each other can move research, and subsequent funding, forward.

START YOUR ENGINES

The first considerations should be made while negotiating a job offer. Candidates should review potential lab space and submit a 'wish list' to the department chair.

"Write down exactly what you think you will need. And I mean everything," says Markus Babst, a cell biologist at University of Utah in Salt Lake City. "If you forget something important, then you will be paying for it yourself later." Don't assume that anything in the department is available to share, he says, or that things such as water-purification systems are in place. Don't worry about going overboard, he adds — the chair will expect to review the list and scale it back as necessary.

To minimize start-up costs, consider the infrastructure, core facilities and expertise that are already in-house, ready to be used, says Elias Arnér, a medical biochemist at the Karolinska Institute in Stockholm. His colleague Marie Wahren-Herlenius,

a rheumatologist, agrees. She points out that this strategy can also align your lab with senior labs that typically attract more good talent than they can take.

Candidates needing their own mass spectrometer or fancy microscope must delicately negotiate it before signing on the dotted line, says Matthias Mann, director of the Center for Experimental BioInformatics at the University of Southern Denmark in Odense. "It's kind of like a courtship situation," he notes, but a delay or misunderstanding could mean years of downtime.

If you need major lab space renovations, these should be included in your job offer. Get over any uneasiness about 'talking money' and asking for things in writing — candidates "should know that's the game", Babst says.

McClure adds that new faculty members should take advantage of the transition time to plan their lab's future. Write a list of everything you will need and try to identify how each piece will be provided. Think about when to submit your first grant proposal and which agency programme will best fit your project. Finally, try to anticipate how to match start-up funds with a first grant budget to maximize spending and minimize funding gaps.

Strategies for spending start-up funds range from conservative to liberal. Hull warns colleagues not to penny-pinch too much. "It's tempting, but it won't get your lab going," she smiles. Babst, who spent three years in the biotech industry, agrees that scrimping doesn't always pay

dividends later. "I learned that time is the key, not money. If it will help in getting results, then I will spend it," he says. After set-up costs, he took the remainder of his start-up funding and calculated how many employees he could pay for two years. "I won't have much left at the end," he admits, but he thinks he will be better positioned to secure a grant.

Both Hull and Babst hired experienced lab technicians, gambling that the investment would pay off in efficiency. Arnér, on the other hand, never hired a technician and says that advising students keeps him closer to the data. Salaries dominate budgets, so investigators should decide first on who they need.

Things that increase efficiency, such as pre-mixed culture media or a better pipette, may be worth the extra money. "A lot of beginners are simply too cheap," says Babst. But with big-ticket items, it is okay to scrimp. It might save thousands of dollars, which can be used to cover unanticipated costs. "Don't buy the Rolls-Royce if you can do it with a Hyundai," he says.

Babst bought freezers from LabTrader, an online



Mike McClure says it is important to plan a new lab's future carefully.

NIH/SH/NIH

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

H. PRINZ/CORBIS

Getting the sums right: make a complete list of everything you need, don't be hesitant to talk about money and ensure you have everything in writing.

company that sells second-hand scientific equipment. He asked colleagues if cheap equipment models worked well or not. Hull struck up discount deals with vendors and scavenged equipment from a retiring lab.

Cost-sharing helps eke out that first major grant budget and sometimes earns points with reviewers. Although the grant-review process has placed less scrutiny on individual budget items recently, there are still some general rules to follow.

"Do not develop a budget without talking to someone who has done it before — ever," says Rebbecca Moen, assistant director of the office of research administration at Duke University School of Medicine in Durham. Get a senior colleague in your field or department to look at the project and decide what you really need, she says. Moen says budgeting from the bottom up will make a budget easier to defend to reviewers.

Some agencies have turned to modular grants in which an investigator applies for a certain number of \$25,000 budget increments. But McClure says it is still not appropriate for new investigators to ask for major equipment. Making unreasonable requests or ignoring programme limits could easily doom a proposal.

"It is critical to know what the granting agency will support and what it won't," says Raelynn Potts, the administrative officer for the Scientific Computing and Imaging Institute at University of Utah. Budget items should be inherently obvious to the benefit of the grant and not just a technical wish-list, she says.

She encourages investigators to nurture relationships with a departmental or university grants officer such as herself, as well as with the granting agency's programme officer, whose job is to answer questions.

First-time grant applicants should try to match

future lab tasks to people. It is difficult, she says, but paying an unexpected salary is near-impossible. And don't forget to factor in sizeable items such as service contracts, employee benefits and tuition for students.

Most departments track researchers' budgets through accounting, but ultimately investigators must monitor spending. "Most researchers want to think about budgets as linear, but the spending on a grant ebbs and flows," says Potts.

Hull says her career award allowed her to take risks in spending to get her lab going and in the research projects she chose. The Burroughs Wellcome programme was designed to provide bridging support to postdocs with a five-year, \$500,000 award that can be carried to their first faculty positions.

McClure says transition or career-development awards give young researchers power to shift money around. "Learn quickly that no research career is supported by a single source," he says. "The cost and size of what you are doing change over time."

Arnér and Wahren-Herlenius note that, in Europe, young investigators usually fund their first two to three years with multiple smaller grants from private sources before applying for larger grants from government research councils. Arnér suggests writing several grants a month in the beginning and Wahren-Herlenius says stay organized with a calendar of deadlines.

McClure suggests arranging large grants so that they overlap in time a bit to avoid any funding gaps. Then, he says, "decorate the base funding with one-year awards" of \$75,000 or so whenever possible. "A proper chief executive will find resources in any amount at any time," he notes. Getting funded is a never-ending process, he says. "But it becomes easier." ■

Kendall Powell is a freelance science writer in Broomfield, Colorado.

**"Don't buy the
Rolls-Royce if you
can do it with a
Hyundai"**
— Markus Babst

GRADUATE JOURNAL

First-night nerves

My first talk at a national meeting provided me with a great educational experience — and a bit of stress. The day before I left for the meeting, my adviser asked me to add some new details. This initially threw me, because I had thought my talk was set — adding more details could push me over my allotted time limit.

Apart from this, I was also worried about the question-and-answer session, because this is where the unexpected can happen. I was concerned that no one would have any questions or that someone would ask something I couldn't answer.

Both those concerns led me to practise and perfect my talk — meaning I didn't have so much time to take detailed notes on other speakers and poster presentations.

In the end, my nerves worked to my advantage. After I cut out some flashy but unnecessary animations, my talk finished on time. My preparation for the question-and-answer period paid off, too. I fielded several great questions about issues that I had thought about and so could address confidently. That felt really good. And focusing on my own talk stopped me from rushing to as many sessions as possible. Instead, I attended only those talks and poster presentations I felt were essential. The whole experience gave me fresh inspiration — and ended without trauma. ■

Anne Margaret Lee is a graduate student at Harvard University.

SCIENTISTS & SOCIETIES

Postdoc problems — postdoc solutions

Since it was set up in 2003, the National Postdoctoral Association (NPA) has been raising awareness about the challenges faced by postdocs in the United States. It advocates three main collaborative approaches to help solve these issues: positive change, the provision of resources to empower stakeholders who seek change, and the formation of a network and virtual gathering place for postdocs and their allies.

Recent months have seen developments in each of these three areas. Advocacy has brought an increase in the amounts that institutions can receive in training grants from the National Institutes of Health (NIH). The NPA will also play an active role in encouraging the NIH to implement the recommendations in the

National Research Council's recent *Bridges to Independence* report.

Meanwhile, the NPA has developed three new resources to improve the postdoc training experience. The *Recommended Practices Guide* targets leaders of research institutions and presents recommendations from national organizations such as the National Academy of Sciences and the Association of American Universities.

The Postdoc Association Toolkit aims to help postdocs start up or maintain a local postdoc association. This online toolkit includes sample by-laws, budgets and guidance on how to represent postdocs within a larger research institution.

And the International Postdoc Survival Guide offers guidance to postdocs from outside the United States. The website explains the differences between a postdoc in the United

States and other countries, provides a comparison of the different types of visa available to foreign postdocs, and gives advice on how to negotiate with a principal investigator.

The third prong of the NPA's work has focused on networking and community-building activities. The association's third annual meeting, held in San Diego last month, brought together postdocs, faculty members and administrators from research institutions and national organizations across the United States. Participants debated the best strategies for seeking positive change.

As the national conversation about the postdoc experience shifts from identifying problems to focusing on solutions, the NPA will continue to work for the best interests of postdocs. ■

Keith Micoli is executive board chair, National Postdoctoral Association. ♦ www.nationalpostdoc.org

MOVERS Tim Berners-Lee, professor of computer science, University of Southampton, UK



One thing Tim Berners-Lee would like, he says, is an in-depth knowledge of a subject area: he calls himself "jack-of-all-trades, master of none". From the inventor of the World-Wide Web, that seems a startling claim.

But Berners-Lee may be on course to get his wish, as he takes up an appointment at the University of Southampton, UK. There he will immerse

himself in the Semantic Web, which aims to give information a clearly defined meaning, so that computers and people can work together more effectively. He will retain his current professorship at the Massachusetts Institute of Technology, but by working more closely with Southampton, he hopes to give the semantic project increased momentum.

Berners-Lee followed an unusual path. The son of two mathematicians, he grew up with a love of electronics. On his way to a first-class degree in physics at Oxford, he spent much of his time building a computer from an old television set and some cheap parts.

"At the time, I was a budding engineer," he says. "I didn't personally know any people doing computer-science research, who might have served as role models. Maybe I should have gone to the United States to do a masters and then come back and done a PhD in computer science."

Turning away from academia, he took a series of jobs in Dorset — largely because he liked the seaside location. In 1989, as a database engineer at CERN, the European Particle Physics Laboratory, he got on with another spare-time project, and the World-Wide Web was born.

"An awful lot of things had to be in alignment for the web to take off," he says. "For the first few years, it didn't have the critical mass to make the world sit up and take notice. I and others at CERN did a great deal of work talking to people and encouraging them to give it a try."

Now, with the Semantic Web, he faces a similar challenge — getting people to adopt a new technology when others are already in place. But he is exploring how the Semantic Web could be useful for life sciences, which he is learning more about on the fly. And in the process, another field may benefit from Berners-Lee's jack-of-all-trades inventiveness. ■

CV **1994–current:** Senior research scientist, Massachusetts Institute of Technology and director, World Wide Web Consortium.

1984–94: Computer engineer, CERN, Geneva, Switzerland.

1981–84: Head of technical design, Image Computer Systems, Wimborne, Dorset, UK.

1980: Consultant software engineer, CERN, Geneva, Switzerland.

1978–80: Software designer, D. G. Nash, Dorset, UK.

1976–78: Principal engineer, Plessey Telecommunications, Dorset, UK.

Oscar night, 2054

A collection of heavenly bodies.

Syne Mitchell

Total Health™: You, Only Better! presents a night at the Oscars. Beatrice Barnard and George Ford help you sort fashion triumph from catwalk disaster. Remember, with *Total Health™*, it's still you — only better!

BB: First on the carpet tonight is pop singer Alexa DuBois, stunning in an off-the-shoulder Giovanni dress whose mocha silk highlights the leopard pattern of her skin. Her body modifications were done by the renowned designer Leonardo Fontesca, so it's no wonder that the work is subtle and flattering.

GF: At her side, sporting a crest of blue-black feathers, is Howe Yorkins, the tennis star. Alexa may look like the cat who swallowed the canary, but he belongs back on the poultry pharm.

Next up is Neo Washington, up for best actor for his part in the ground-breaking third remake of *Gone with the Wind*. Off the set, he flaunts his political vegetarian stance with a photosynthesizing teal complexion. The Greens behind the velvet ropes have gone insane, cheering and holding up 'Meat Kills' signs. Only Neo could mix ethics and fashion and make it work. We love you, Neo!

BB: Following Neo is the quartet of AI specialists whose virtual personality, Babylon Brown, is nominated for best actress for her role as Hillary Clinton in the historical drama *You Go, Girl*. They've made an attempt at cohesion, if not style, by wearing identical white lab coats over jeans and black T-shirts, bought online from AmazonMart, no doubt. Their pedestrian modifications are typical gamer-geek fare: enlarged eyes, polydactyl digits, and a revved-up nervous system that makes them twitch like crickets on a griddle.

GF: Back to the lab, boys!

On the carpet now is Medusa Addams, the character actress we all love to hate. Her puff-adder up-do is a-glitter with diamonds as each of her 24 snakes is wearing a custom-fit tiara. She completes the look with a skin-tight snakeskin sheath — oh wait! That is her skin. Cheeky, Medusa, very cheeky.

BB: Here's fitness guru Radha Guriampadaya, wearing an uninspired electric-blue bikini over buzz-cut fuchsia fur. Hell-o! Anime is so O-V-E-R!

GF: Her companion, Susanna Belle, is positively fey in a bell-shaped diaphanous silk shift, complemented by feathery antennae and glistening dragonfly wings.

Rumour around town is that her geneticist has her moulting a new pair every 30 days, in a painful three-day process...

BB: Oh — my — *GOD*. George, you'll have to take this one — I'm speechless.

GF: It's ageing megastar, Monica Moe, and what a display she's putting on tonight! If you're not getting a video or threed broadcast, you must check out the Internet archives. Sporting what can only be offshore modifications, she's got cleavage down to her hips, and I mean that literally. In white pleated chiffon, she looks like a cross between Marilyn Monroe and Artemis Ephesia. I count four, six, no — eight pairs of breasts. Well, slap my cheeks and put me on the cover of 'Trying Too Hard' monthly.

BB: Just stepping out of the limo is teen superstar Alessandro Goldstein. The crowd's gone wild. He — yes, it's true — the child prodigy singer/actor/supermodel is classic and divine in black leather pants and jacket, gold-net undershirt and sporting an amazing retro-humanist look with no obvious modifications. Could that flawless body be natural? Only his geneticist knows for sure.

GF: At Alessandro's side is his child bride and the youngest Nobel laureate, Dame Bekka Davis Llewellyn, who won last year's highest scientific honours for her work on inserting and upgrading artificial human chromosomes. She positively glows in a cowl-neck gold-lamé top over an A-line floor-length velvet skirt. Very Audrey Hepburn. Like her husband, she displays no obvious enhancements ... although nasty rumours

after last year's Nobel awards implied that her intelligence may not be quite so native.

BB: Between them, they exemplify the new post-post-human movement in fashion. It's regrettable that such extremists have found a place in the public eye. The effect they are having on today's youth is deplorable. Anywhere you look you can see young people flaunting genetic flaws: buck teeth on the Transatlantic Tunnel, stringy lifeless hair on the Prada Channel, even love handles at the Milan Show for fashion's sake!

GF: Don't worry Bea, retro-humanism is a passing fad. Body modifications will always be with us. Engineers will always have work. Gene-Es don't go back in bottles.

Total Health™ reminds its subscribers that due to a shortage of recombinant DNA reagents, this year's chromosomal updates are only available to those subscribers who are critically ill, under 24 months in age, or whose updates are more than two years out of date.

Syne Mitchell is an award-winning author of several science-fiction novels, including biotechnology thriller *The Changeling Plague* and near-future space adventure *End in Fire*. She lives in the foothills east of Seattle with her husband, SF-writer Eric S. Nylund, and their toddler son Kai. She is resplendent in authorial post-fashion irony: denim overalls over a hemp shirt and antique Birkenstocks. No body modifications are visible.



JACEY